STATISTICS NETHERLANDS

Usual Residence Population Definition: Feasibility Study The Netherlands

Contacts:

Mr. Wim Leunis
Head of Demographic Statistics
Statistics Netherlands
P.O. Box 24500
2490 HA Den Haag
Nederland
E-mail: w.leunis@cbs.nl

CBS Den Haag
Henri Faasdreef 312
2492 JP The Hague
P.O. Box 24500
2490 HA The Hague
+31 70 337 38 00

www.cbs.nl

project number 301195; 300634
BPM; SES
30 December 2016

Index

PREFACE

This report presents the Dutch feasibility study regarding the possibilities in the Netherlands to make the national and regional population statistics available according to the concept of usual residence in the European Demographic regulations. This feasibility study was requiredin article 8 of the Regulation (EU) No 1260/2013 of the European Parliament and of the Council of 20 November 2013 on European demographic statistics and in the Draft Grant agreement 07121.2014.006-2014.357 signed for the Commission on 23 September 2014 by Mr. Gallo Gueye, Director, Directorate F on 23 September 2014 and for Statistics Netherlands by Mr. Tjark Tjin-A-Tsoi, Director General on 21 October 2014.

The report is structured according to the document 'Usual Residence Population Definition: Feasibility Studies'. In this report the chapters are set up according to that document.

As a first step in preparing the study underlying report useful and highly appreciated contributions were made by Prof. Dr. G.B.M. Engbersen (Rotterdam Erasmus University), Prof. Dr. P.G.M. van der Heijden (Utrecht University) and Dr. R.P.W. Jennissen (Ministry of Security and Justice Research and Documentation Centre (WODC)/Netherlands Scientific Council for Government Policy (WRR). However, Statistics Netherlands is solely responsible for the views and results presented in this report.

It is kindly requested to send any communication about this report to the following address:

Mr. Wim Leunis
Head of Demographic Statistics
Statistics Netherlands
P.O. Box 24500
2490 HA Den Haag
Nederland
E-mail: w.leunis@cbs.nl

SUMMARY

Statistics Netherlands has applied two different methods to estimate the usual residence population:

One method was based on applying a capture-recapture (CRC) method to the combination of three registers. The general idea behind the CRC method is based on the situation of having two consecutive samples from a single population. The "captured" sampling units in the first sample are marked and set back into the field. Then the second sample is "captured" and the number of "newly" captured as well as the number of recaptured (i.e., marked) sampling units is counted. Based on those counts, an estimate is constructed for the total population. Applying this to registers instead of samples, an estimator can be constructed for the under coverage of each of the registers. This was extended to the situation with more than two samples/registers. However, this method is not fit for estimating over coverage.

The other method (micro register data method) was operated at the level of individual records. The essence of this method is that records are explicitly added to or deleted from the population register data. Micro data from various registers are combined in order to obtain estimates of the usual residence population at the national and the regional level. Contrary to the former method, no mathematical estimation model was followed. The method merely combines data. Since data sets about a number of persons are lacking, this method leads to underestimates of both under coverage (such as persons whose stay in the country is illegal) and over coverage (such as persons who left for a foreign country and did not notify the population register).

Consequently, both methods have advantages and disadvantages. The capture-recapture method gives an interval estimate of the under coverage, but this method is not fit for estimating the over coverage. The micro register data method produced both under coverage and over coverage figures, but both are underestimated due to missing registers about a number of groups, such as illegal or undocumented persons.

Combination of the results from the two methods has led to the conclusion that the usual residence population on 1 January 2013 is 16,916.3 thousand, which is 0.8 per cent higher than the official population size, derived from the Dutch population registers. If it is assumed that the ratio between these two figures remain constant in 2013-2016, the following results are obtained:

Usual residence population, 1 January 2013-2016

	Population size under national concept	Usual residence population
	x1,000	
2013	16,779.6	16,916.3
2014	16,829.3	16,966.4
2015	16,900.7	17,038.5
2016	16,979.1	17,117.5

The 1 January 2013 usual residence population size appears to be 136.7 thousand higher than the population number derived from the population register. The under coverage estimate is 169.9 thousand, the over coverage estimate is 33.2 thousand.

The figures on the Dutch usual residence population published by Eurostat so far (2014-2016) were based on an extrapolation of the 2010 estimate, assuming equal developments for registered and non-registered usual residents (Annex 5). There are two big differences with the estimates presented in this report:

- In the present estimates under coverage and over coverage were taken into consideration. The former estimates presented only under coverage.
- Since 2014, the capture-recapture method has been improved in a number of steps. The methodology, the software and the estimation of the under coverage of the 0-14 years old were improved, thereby improving compliance with the assumptions underlying the CRC method. The figures presented in Gerritse et. al (2016) were based on a number of these improvements. In the current report the method was improved further. As a result the present width of the uncertainty interval surrounding the estimate is now much smaller than the 2010 one.

Today the original 2010 estimate is to be considered as rather rudimentary. Starting from 2017, the annual update of the usual residence population size will be based on the current method. Obviously, this will lead to a break in the time series already provided and we propose to discuss with Eurostat how to deal with this.

The decreased interval width gives Statistics Netherlands more confidence in the new estimate than in the former one. However, given the uncertainties attached to the Crime Suspect Register used in the capture-recapture method, the missing data for some population groups and the assumptions that needed to be made when applying the CRC method and the micro register data method it must be stressed that neither the former estimate nor the new one can be considered to produce a definite reliable estimate of the usual residence population size. For this reason Statistics Netherlands reserves the right to revise the estimates when new methods and techniques or new data sources become available.

Since application of the two methods and combining their results appears to be very time consuming, it is impossible at the moment to carry out this procedure on an annual basis. For reasons of validation it should, maybe partially, be repeated after a

number of years. A cost-benefit analysis will be carried out in order to establish the most appropriate approach for the years to come.

Application of the capture-recapture method has only led to figures at the national level. This method could not produce data at the regional level. This is why data at the regional level were arrived at by application of the micro register data method, by retrieving the region were the residents were registered on 1 January 2013. Unfortunately, this regional information was not available for all residents. There were 158.1 thousand unknowns in total. This corresponds with 0.9% of the population. Obviously all unknowns are part of the under coverage (93%).

Regarding vital events no numbers of births and deaths to usual residents are available. However, the birth and death certificates issued by the local registrars give an indication that no births are missing and the number of missing deaths appears to be minimal.

1. National and regional demographic statistics

The Dutch demographic statistics are entirely based on the Dutch population registers. As such, describing demographic statistics in the Netherlands basically boils down to describing the definitions and practices used in the population registers. In this document these registers are referred to as PR.

The PR Act sets rules for registration and deregistration. Everyone who enters the Netherlands is registered as a resident (immigrant) provided

1. his/her stay is legal according to the Immigration Act (on people who do not have the Dutch nationality);
2. the intended stay is at least two thirds of the forthcoming six months;
3. the person is properly identified. The latter means that a valid passport or other official document is shown for identification.

Every child born in the Netherlands whose mother is registered as a resident is also registered as a resident. Children who are born abroad to a mother who herself is registered as a resident of the Netherlands are registered (as immigrants), provided the children will live in the country.

There are two exceptions to the rules for registration. The first applies to so-called privileged persons, including foreigners working on Dutch soil as diplomatic, consular or military officials, or in an international organisation. Since they enjoy a special 'privileged' status and are not considered foreigners under the auspices of the Immigration Act, they are given the choice whether or not to be entered in the population register. Asylum seekers form another exception. Their registration takes place only six months after their arrival in the Netherlands, irrespective of their expected duration of stay, unless they are granted a residence permit within six months. In that case they are registered when the residence permit is granted. However, children born to asylum seekers who are not yet registered are registered directly after birth. This leads to the somewhat odd situation that a new-born baby is registered whereas the parents and siblings are not (yet).

Deregistration as an inhabitant takes place when someone dies, or leaves the country and intends to stay abroad for at least two thirds of the forthcoming twelve months (emigration). The rules concerning registration and deregistration make no distinction between Dutch and non-Dutch nationals. Of course, the Immigration Act is not applicable to Dutch nationals.

Annex 2 gives a more extensive description of the definitions and practices used in the population registers.

Basing the Dutch demographic statistics on the population registers implies that every person who is registered as a resident is included in the demographic statistics. No resident is left out, and no other person is included. Once a year, around 1 February, Statistics Netherlands obtains data about all persons who were ever included in the population register in 1995 or later. Data about births, deaths, changes of address (including immigration and emigration), marriage etcetera, are obtained on a daily basis. This enables Statistics Netherlands to produce the number of inhabitants for every reference date of the calendar year. In practice, data are produced on a monthly

basis, viz. 1 January, 1 February, etcetera. Vital statistics are produced per calendar month and calendar year. Moreover, mortality figures are produced at a weekly basis.

For each of the 21 groups mentioned in the EU document, the inclusion or exclusion in the demographic statistics is described in Annex 1.

Vital events

Births are declared at the local registrar of the municipality where the child is born. However, the child is entered in the population register of the municipality where the child lives. Usually this is the municipality where the mother is registered as a resident. Still-births are not registered in the population register. However, the local registrars report them separately to Statistics Netherlands.

Like all demographic statistics, birth statistics are fully based on the population registers. All live births are included in the birth statistics. However, children born to asylum seekers who are not (yet) registered in the population register, form an exception. When registered (six months after their arrival in the Netherlands) these parents and siblings are obviously included in the immigration statistics. This is why Statistics Netherlands decided to include also the newly born in these immigration statistics (instead of the birth statistics).

Children born to mothers who are not registered as residents themselves are not included in the birth statistics or any other statistics related to the population of the Netherlands.

Another category that needs attention are children born abroad to mothers who are registered as inhabitants of the Netherlands. Those children are included in the Dutch birth statistics only if they come to live in the Netherlands within seven days after birth. They are not included in the immigration statistics, but after eight days or later they are.

All deaths registered in the population registers are also included in the death statistics, including residents who died abroad. People who are not included in the population register who die while in the Netherlands are not entered in the Dutch death statistics.

Immigration and emigration

Everyone who enters the Netherlands is registered as an resident (immigrant) provided

1. His/her stay is legal according to the Immigration Act on people who do not have the Dutch nationality;
2. the expected stay is at least two thirds of the forthcoming six months;
3. the person is properly identified. The latter means that a valid passport or other official paper is shown for identification.

Emigration relates to persons who leave the Netherlands for at least two thirds of the forthcoming twelve months and who inform the municipal authorities of their departures. They are included in the Dutch emigration statistics. Unfortunately, about

one in three persons who leave the country do not notify the municipal authorities of their departure. When the authorities find out that someone is 'missing', the law stipulates that they must investigate his or her whereabouts. In Annex 2 a detailed description is given of what steps must be taken. If the investigations lead to the conclusion that the persons remains missing, he or she is administratively emigrated to a country qualified as 'unknown'. They are included in the statistics of administrative removals. If they return to the Netherlands they are included in the statistics of administrative entrances.

Demographic balancing equation
Naturally, for each population the following equation holds:
$P_{t+1} = P_t + B(t, t+1) – D(t, t+1) + Arr(t, t+1) – Dep(t, t+1)$
where

P_t = the population size at date t

B(t, t+1) = the number of births between t and t+1

D(t, t+1) = the number of deaths between t and t+1

Arr(t, t+1) = the number of arrivals between t and t+1

Dep(t, t+1) = the number of departures between t and t+1

Given the definitions of immigration, emigration, administrative entrances and administrative removals in the Dutch demographic statistics, and restricting ourselves to calendar years, the demographic balancing equation can be rewritten as follows:

$P_{t+1} = P_t + B(t) – D(t) + Imm(t) – Em(t) + AdmEnt(t) – AdmRem(t) + Corr(t)$

where

P_t = the population size at the beginning of calendar year t

B(t) = the number of births during the calendar year t

D(t) = the number of deaths during the calendar year t

Imm(t) = the number of immigrants during the calendar year t

Em(t) = the number of emigrants during the calendar year t

AdmEnt(t) = the number of administrative entrances during the calendar year t

AdmRem(t) = the number of administrative removals during the calendar year t

Corr(t) = the number of corrections needed to have a balanced equation

The population size at the beginning of a calendar year is mainly based on the Annual Selection data obtained from PR around 1 February of that calendar year. The term Corr(t) is needed because in practice persons may appear in or disappear from the Annual Selection who cannot be attributed to births, deaths, immigration, emigration, administrative entrances or administrative removals. Fortunately, their number is limited. In 2010-2015 the annual average amounted to 0.4 thousand.

2. Gaps between national and usual residence population definitions

Following Regulation 1260/2013 'usual residence' means the place where a person normally spends the daily period of rest, regardless of temporary absences for purposes of recreation, holidays, visits to friends and relatives, business, medical treatment or religious pilgrimage. The following persons alone shall be considered to be usual residents of a specific geographical area:

(i) those who have lived in their place of usual residence for a continuous period of at least 12 months before the reference time; or
(ii) those who arrived in their place of usual residence during the 12 months before the reference time with the intention of staying there for at least one year.

There are a number of gaps between this definition and the national definition of the Netherlands. According to the latter definition, the intended duration of stay is much shorter than in the usual residence definition, to begin with. This means that a correction is needed for the number of persons whose stay in the Netherlands is less than twelve months. With respect to emigration, people who leave the country for more than eight months are excluded from the population according to the national definition. It follows that if people stay abroad for less than twelve months this must be taken into account given the usual residence definition.

The national definition concerns people whose stay in the Netherlands is legitimate. Persons whose stay is illegal according to the Immigration Act are not registered in the population register, they are not included in the national population figures either. Seeing that the usual residence population definition does not exclude people whose stay is illegal, they must be added to the latter population.

Contrary to the usual residence definition, the national definition has several exceptions on the registration of various population groups. Privileged persons are given the choice whether or not they wish to be registered. About one in three decide to do so, which means that the other two in three are not included in the population figures according to the national definition. This group includes people who are working as foreign diplomatic, consular or military officials, or for an international organisation, and their family members.

Asylum seekers are the other exception as to being registered in the population register. Their registration takes place six months after their arrival in the Netherlands, irrespective of their intended duration of stay. It is commonly accepted that asylum seekers generally have the intention to stay in the country for at least twelve months. From this it follows that they should be included in the usual residence population from the day of their arrival. However, restricting the inclusion to people who are granted a residence permit is considered to be a better approach, seeing that not everyone is allowed to remain in the country. Another consideration at this point is that a number may decide to leave the country for another EU member state.

As pointed out in the previous chapter, the Dutch population registers are kept by the municipal authorities, who must apply the rules laid down in the PR Act. However, it is acknowledged that in some cases it is difficult for municipalities to keep to these rules.

Such difficulties are usually the result of persons not informing the municipal authorities of their arrival, departure or move elsewhere within the country. Annex 2 elaborates more extensively on how the municipal authorities deal with such cases.

Incorrect registrations have a direct influence on the content and the quality of the demographic statistics. Persons who are registered after a delay, such as asylum seekers or late registrations, are not covered by these statistics during the first weeks or months of their stay in the Netherlands. Other categories, such as foreign diplomats or non-Dutch NATO military, may not be covered at all. On the other hand, persons who leave the Netherlands without informing the municipal authorities of their departure, are inevitably covered during the weeks and months when they have no actual residence in the Netherlands.

In the next chapter it is proposed how to fill in these gaps.

3. Proposal of how to fill in the gaps; methods description

The number of usual residents at the national level is estimated in two ways. The results from both methods were compared and evaluated and one best figure was decided upon.

Applying two approaches is considered appropriate given the fact that no data sources are available from which the usual residence population can be directly and fully derived. It was hoped that if two distinct methods lead to more or less the same results, one can be more confident about the reliability of the results.

3.1 Method 1[1]: Capture-recapture estimation approach

This method is based on applying a capture-recapture (CRC) method to the combination of three registers: the (Basic) Population Register (PR), the Employment register (ER) and the Crime Suspects Register (CSR). The PR is the backbone of many social surveys conducted by Statistics Netherlands, the ER contains information on all employees in The Netherlands and the CSR contains information on all persons reported to the Dutch police force as crime suspects.
The general idea behind the CRC method is based on the situation of having two consecutive samples from a single population. The "captured" sampling units in the first sample are marked and set back into the field. Then the second sample is "captured" and the number of "newly" captured as well as the number of recaptured (i.e., marked) sampling units is counted. Based on those counts, an estimate is constructed for the total population. Applying this to registers instead of samples, this means that an estimator can be constructed for the under coverage of each of the registers. This can be extended to the situation with more than two samples/registers, see e.g., Fienberg (1972), Bishop *et al.* (1975) and Van der Heijden *et al.* (2012).
Applying CRC, the following is assumed:

1. Closed population (no units enter or leave the population)
2. Independence of inclusion probabilities of the different sources (probability of being included in the second register is independent from the probability of being included in the other register)
3. No erroneous captures (no capture of units that do not belong to the target population)
4. No errors in linking records between the different sources

In practice however, these assumptions cannot always be met, particularly when data are originally not collected for statistical purposes.
Gerritse *et al.* (2016) have applied the CRC method to estimate the under coverage of the PR, ultimo September 2010, for the ages 15-65. We have adjusted that methodology slightly and produced an estimate of the under coverage of the PR ultimo March 2014, for the ages of 15-65. The fact that we estimate the under coverage with

[1] The description of this method is based on the paper presented at the Conference of European Statistics Stakeholders (CESS) meeting, 20-10-2016, Budapest, Hungary; Bakker and Prins (2016).

CRC only for the ages 15-65, is due to the fact that the ER is only applicable to that age-group. We will estimate the under coverage of ages 0-14 and 65+ by other means.

Dealing with the CRC assumptions
The assumption that the population should be closed, is easily met by the PR and the ER. Both registers describe a period and any common date or period can be selected. The assumption is satisfied by restricting the data to the status of only one day: *ultimo* March 2014. For the CSR however, is event-based: crime suspects are registered of which the police make a report. The number of events on one specific day is not enough to apply the CRC method. In order to satisfy the assumption as well as possible, we have restricted the period of the CSR to the first half year of 2014.
The assumption on the independence of the inclusion probabilities can be relaxed by using three registers: the inclusion probabilities between any two registers should be independent, conditionally on the third register. Moreover, we have used three covariates in order to further relax the assumption on independence of the inclusion probabilities: nationality, age-classes and gender. We now only assume that the three factor interaction between the inclusion probabilities of the three registers is zero within each cell of the cross table of the covariates.
The assumption on erroneous captures is taken care of by removing units from the PR, ER and CSR of persons that to not belong to the population:

a) The few persons with Dutch nationality not registered in the PR (we expect them to be expats working in another country)
b) Persons with an address in Germany or Belgium (it is likely they live in Germany or Belgium and only travel to the Netherlands to work, go to school, shop or have a short holiday)
c) Persons who are reported for a crime by the border police (they would not be allowed to enter the Netherlands)

However, after removing those units, it is still possible that the CSR contains persons who committed a crime but do not belong to the Dutch population. E.g., drugs runners, gangs of pickpockets, tourists arrested for drunk driving. We dealt with that by applying different scenarios to deal with the records in the CSR that are not linked to the PR nor the ER. We assumed different proportions of those records to be either erroneous captures (and would thus not be used in the CRC estimate) or missed links (and thus would be added to the number of links with PR and/or ER). Calculating the CRC estimate under these scenarios gives some information on the sensitivity of the method to the presence of erroneous captures.
The assumption of no linkage error is not completely met, despite the careful record linkage procedure we used, as described in Annex 3. Two types of linkage error may occur: miss-links (linking two records that should not be linked) and missed linked (not linking records that should be linked).
We assumed that the linkage errors only occurred in the probabilistic record linkage step. The probabilistic record linkage step provides some information on the probability of the occurrence of those errors. In Tuoto (2015) an adjustment of the standard CRC estimate is given, to take into account the linkage errors in case of linking two registers. To be able to use their method, the probabilities for occurrence of the linkage errors need to be estimated quite accurately. In our situation this turned out to be too complicated, due to blocking, the use of covariates and the fact that we actually

use three registers. Using simulations we found that using not too accurate estimates of the error probabilities in the corrected estimator, could yield worse results compared to the non-corrected estimator. Therefore, we decided not to apply this correction to our CRC estimate.

Estimation of under coverage of 0-14 year-olds
To estimate the under coverage of the PR with respect to 0-14 year-olds, we made use of a register on enrolment in education. The assumption is that all children, whether registered in the PR or not, must attend (primary) school. Children must enter (primary) school for the first time in the year they turn 5. Using this register on enrolment in education, we could count the number of children aged 5-14 that are present, or not, in the PR. We used the enrolment registers for 1 October 2012 (school year 2012-2013), 2013 (school year 2013-2014) and 2014 (school year 2014-2015). Using those three registers, we could derive whether or not the children were enrolled in education for one year on 1 October 2013 and 1 October 2014. Averaging over these two dates yielded our estimate for under coverage of 5-14 year-olds ultimo March 2014.
Using the estimate for the ages 5-14, we extrapolated the results to the ages 0-4, by assigning the same under coverage of the PR (as an absolute number) from age 5 to each of the five ages 0-4.
We ended up with the following estimate of the under coverage of the PR for persons 0-14 years of age (multiples of 1,000):

	1-10-2013	1-10-2014	Ultimo March 2014
0-14 years old	7.6	7.2	7.4

Estimation of under coverage of persons 65+
We did not have any register information outside the PR available for the people aged over 65. Therefore we had to make an educated guess. Qualitative research hinted that this number should be limited to a few thousands. In Gerritse *et al.* (2016) it was argued that estimating their numbers using the PR and the CSR would yield a number between 1,000 and 10,000.

Differences in implementation, compared to the method in Gerritse et al. *(2016).*
There are some small differences in the implementation of our estimation procedure with respect to the method used in Gerritse *et al.* (2016):

- In our approach we selected persons with duration of stay of at least one year before applying the estimation procedure. In Gerritse *et al.* (2016) the duration of stay was used as an additional covariate in the estimation procedure.
- In our approach we first calculated the duration of stay for all persons in ER and then made a selection ultimo March 2014. In Gerritse *et al.* (2016), first a selection of records ultimo September 2010 was made and then the duration of stay was derived.
- We imputed duration of stay in CSR whenever missing, prior to the estimation procedure (see Annex 4). In Gerritse *et al.* (2016), this was part of the estimation procedure itself.

Final estimate of the under coverage, using the CRC method

According to the more detailed information in Annexes 3 and 4, we should be aware that the estimate has quite some uncertainty attached to it. Firstly, because the CRC method itself introduces a variance component. Secondly, because the different scenarios to deal with erroneous captures in the CSR show some variation. Thirdly, because the extrapolation of the information on 5-14 year-olds to the 0-4 years-olds is based on an unverifiable assumption that those age classes behave the same with respect to enrolment in education for the children not registered in the PR. And fourthly, because the estimate of the elderly (65+) is an educated guess at best and hence has a large uncertainty attached.

Collecting the different estimates, with a quantification of the uncertainty as far as available, we end up with the following table (multiples of 1,000):

	15-65	<15	>65	Total
Not in PR, but in ER and/or CSR	32.3	-	0.2	32.5
CRC estimate of under coverage[a]	[113.0, 136.0]	-	-	[113.0, 136.0]
Education enrolment	-	7.4	-	7.4
Educated guess[b]	-	-	[1, 10]	[1, 10]
Total				[153.9, 185.9]

[a] lower and upper estimate based on scenarios mentioned in Annex 4.

[b] lower and upper estimate based on educated guesses

Comparison with earlier released estimates

In Gerritse *et al.* (2016)[2] an interval for the CRC estimate of under coverage (in multiples of 1,000) of the age group 15-65 for ultimo September 2010 was given as [54, 151].

The current CRC estimate of under coverage (in multiples of 1,000) for ultimo March 2014, shows a smaller interval for the age group 15-65: [113, 136]. This is due to the fact that we have improved the estimation procedure and that we took some other decisions on the selection of persons in the different registers (see "*Differences in implementation, compared to the method in Gerritse* et al. *(2016).*" above). The smaller interval gives us more confidence in the new estimate.

Obviously, the current CRC estimate also differs from the extrapolated CRC estimate based on 2010 data, because we now used more recent sources (2014 data).

Future estimates

In the CRC method we used three registers: PR, ER and CSR. It turned out that the quality of the PR and ER appeared to be quite stable. This is largely due to the fact that these registers are widely used among several governmental agencies for administrative purposes. By contrast, the use of CSR is limited to police purposes. This may explain the unstable quality of the variables in the CSR that we use as linkage keys in our estimation procedure.

This means that newer estimates could deviate from the estimate as given in this report. Firstly because the quality of the CSR may fluctuate and secondly because we might decide to use a different, and preferably more stable, third register in our CRC method in the future.

[2] In that paper a part of the methodological improvements established since 2010 were applied. See section 3.4..

3.2 Method 2: Micro register data approach

The other method is operated at the level of individual records, it is referred to as micro register data method. The essence of this method is that records are explicitly added to or deleted from the population register data. Data from various registers are combined in order to obtain estimates of the usual residence population at the national and the regional level. Given the high quality of these registers this approach was considered promising and effective. Indeed, next to the population register a big variety of activities concerning labour, education, health care etcetera are registered at the individual level. As a rule these register data contain unique personal identification numbers, enabling high quality linking of these data. Unfortunately, some persons residing in the Netherlands, such as illegal and undocumented persons, are not covered by any register. As a result, the micro register data method leads to high quality results for registered persons, but fails to cover all usual residents.

Contrary to method 1, no mathematical estimation model was followed. The method merely combines data, and nothing more than that. However, more registers were involved than in method 1. Since the population register aims at describing the population of the Netherlands, it is considered to be the logical starting point for estimating the usual residence population. As pointed out in the previous chapter, it is acknowledged that the criteria for being registered or deregistered do not fully comply with the usual residence population definition, so information from other sources is needed to obtain better results.

As pointed out in Annex 2, a Citizen Service Number BSN is assigned to each person who is registered in the population register, whether as a resident or a non-resident. The BSN serves as the most important linkage key in governmental and many non-governmental registers, facilitating linkage of registers from a wide range of domains, such as taxes, labour, social security, education and health care. What is more, the linkage quality is extremely high since the uniqueness of BSN is guaranteed. Thus linking population register data to labour, education and other data merely boils down to deciding under what conditions linked data refer to usual residents. In other words, a series of assumptions is needed that tell us which persons are usual residents.

For some groups mentioned in the 'Usual Residence Population Definition: Feasibility Studies' document there is detailed information available, for others groups there is not. To give an example, register data is lacking about most people whose stay in the Netherlands is illegal. People who know that their presence in the country is against Dutch law, will do their utmost to keep out of sight of the authorities, trying to make a living, and stay out of the hands of the police. A complete overview of these persons with information about when they entered the country is an illusion. There are some data about persons who the police came across and who were forced to leave the country. However, in these data not only the duration of stay in the country is missing, the data do not give a clue about the incompleteness either. Van der Heijden et al. (2012) estimate that no more than 7 per cent is caught by the police, so only a fraction is forced to leave the country. What is more, given the limited police capacity, persons who do not make trouble stand a fair chance of being left alone by the authorities. The number of persons whose stay is illegal, is estimated by them as 35.5 thousand (Van der Heijden et al., 2012).

Students in higher education studying abroad is another group about whom there is no full and detailed information. Here a more general issue is touched on. The number of data that can be helpful when estimating the under coverage is bigger than the data used for estimating the over coverage. Dutch data about foreign students at Dutch universities are widely available, whereas at Statistics Netherlands only have limited data are available about Dutch students at universities abroad, if any.

Register data involved in method 2 should give a clue about the duration of stay of each individual in the country. In register terms, it means that a clue is given on the day when people entered the country and when they left again. There is even a possibility that a number of people who entered the country actually never left. The number of registers with such data is scarce. The registers that we had at our disposal provide data about when a job or an education was started or ended. Since no other data were available it was assumed that entering and leaving the country, respectively, took place on the same dates..

Method and assumptions

Some data needed for estimating the usual residence population were already available at Statistics Netherlands, others were not. The Ministry of Foreign Affairs was asked to provide data on foreign diplomatic and military personnel in the Netherlands, and on Dutch diplomatic and military personnel with a post abroad. Data on asylum seekers and procedures were asked for from the Immigration and Naturalization Service. It was decided to restrict ourselves to calendar years for which complete data sets were available or could be provided. Given the fact that such data could be obtained for the calendar year 2013 and earlier, it was decided to estimate the usual residence population as per 1 January 2013. At present the micro register data method does not allow to produce estimates for 2014 and subsequent years. From the usual residence population definition it follows that any person who is in any of these records for twelve consecutive months between 1 January 2012 and 1 January 2014, is in the usual residence population of 1 January 2013. Unfortunately, within the given timeframe it was impossible to arrive at estimates for more recent calendar years as well. However, this does not necessarily mean that it will prove impossible to provide recent estimates in future. On the contrary, the feasibility study shows that some data are more useful than others. What is more, obtaining data from the Ministry of Foreign Affairs and the Immigration and Naturalization Service took much time because these activities are not a part of the regular data collection programme[3]. In chapters 4 and 5 this issue is discussed in more detail.

Method 2 consists of combining the data from all registers (see below), where the date of entry and departure was established for each person. Merging the dates of entry and departure led to a complete overview of when in 2012-2013 a person could be considered to be on Dutch soil. To give an example, if someone who was registered in the population register in March 2013 was still there by the end of 2013, and from the asylum seekers data provided by Immigration and Naturalization Service it turned out that this person applied for asylum in August 2012, the conclusion is that this person is in the usual residence population of 1 January 2013.

[3] Of course it will be possible to have structural data supply arrangements in future in order to ensure timely data supplies.

In practice the method needed to be more sophisticated. Not everyone in the register data was considered a usual resident. Someone who lives in the German town of Aachen, say, who commutes daily his or her job in the Dutch town of Heerlen, is permanently registered as a worker in the Employment Register, but he or she is no usual resident of the Netherlands. Here a number of assumptions were made, where in addition some variants were distinguished, as follows:

- Workers and students whose place of residence is registered as Belgium or Germany (in the registers listed below) are considered to be residents of those countries. When the place of residence is registered to be abroad, but not Belgium or Germany, it is assumed that the time between date of entry and date of departure from the register is fully spent in the Netherlands.
- A non-resident whose place of residence is not Belgium or Germany, who is registered as having two not-adjacent jobs in the Netherlands and the number of days between the two jobs does not exceed 31, is assumed to spend the days between jobs in the Netherlands.
- Residents on 31 December 2012 who died in 2013 are considered usual residents even if the total number of months spent in the Netherlands was less than 12.

Next to the population register data, the following data were used for estimating the usual residence population. The aim was to identify data to be added to or deleted from the population register data in order to arrive at an adequate estimate of the usual residence population.

Data on labour and social security: Employment Register

Every time when salaries are paid (usually once a month) employers make a declaration to the tax authorities. Data about employments and individual labour costs are registered in Employment Register (ER).[4] This register is kept by the Dutch Institute for Employee Benefits Schemes. Individual labour costs data determine the allowance in case of illness, unemployment or inactivity. Once a month Statistics Netherlands is supplied with individual data from ER, needed for the labour and labour costs statistics.

In the feasibility study the following data were used:

- Citizen service number BSN
- Birth date
- Gender
- Address (in the Netherlands or abroad)
- Period when payments for performed labour were received

Data on privileged persons: PROBAS personal records database

The Protocol Department of the Ministry of Foreign Affairs registers all staff members of international organisations and family members forming part of their household in the PROBAS personal records database and, in accordance with the headquarters agreements, provides them with an identity card.

[4] Data from this register were also used in method 1.

For privileged persons registration in the population register is not mandatory but may be done on a voluntary basis. However, there are certain benefits when registering in this register. For instance, the Dutch government only needs to ask for personal information once and the delivery of government services will generally be faster. These benefits are also valid for PROBAS. About one in three choose to be registered in the population register.

Privileged persons obtain a BSN by registration in the PROBAS personal records database of the Ministry of Foreign Affairs. Once a privileged person has been registered, the Ministry of Interior Affairs and Kingdom Relations will automatically issue a BSN. As such, all privileged persons, whether resident or non-resident, are uniquely identified by BSN, in other words: by the same type of personal identification number. This enables Statistics Netherlands to link data from the PROBAS register to other register data, it also avoids double counting, e.g. privileged persons who have registered in the population register.

Registration in the PROBAS register must take place as soon as possible. Therefore, within eight days after arriving in the Netherlands the international organisation should register the privileged staff members and family members forming part of their household with the Protocol Department of the Ministry of Foreign Affairs. What is more, the international organisation must notify the Ministry of Foreign Affairs within two weeks of any changes that occur in the staff members employment or personal circumstances as were recorded in PROBAS. International organisations are responsible for ensuring that the obligation to notify the Ministry of Foreign Affairs is complied with. This notification procedure should also be followed in the event of marriage, divorce or registered partnership of a staff member, at birth of a child, at death and when a change of address occurs.

In order to carry out the feasibility study Statistics Netherlands obtained the following data from the PROBAS registration:

- Citizen service number BSN
- Birth date
- Address
- Date of registration in PROBAS
- Date of deregistration from PROBAS (empty if not applicable)
- Nationality

Data on asylum seekers

As pointed out in chapter 1, asylum seekers are registered in the population register only six months after their arrival in the Netherlands, or within six months after their arrival if they are granted a residence permit. The population register data do not always show who is an asylum seeker. This is why data were obtained from the Immigration and Naturalization Service on persons who applied for asylum in 2012 and persons whose application for asylum was not yet decided upon by 1 January 2013. These data were:

- Citizen service number BSN (if in the population register by the time the data file was compiled)

- Birth date
- Date of application for asylum
- Nationality

Data on pupils and students who pursue their education and studies in the Netherlands

Data about pupils following education that is financed by the government is kept by the Dutch Education Service (DUO). This organization finances pupils and schools and takes care of exams. As such, DUO enables pupils to follow adequate education. DUO does not keep data about pupils at private educational institutes. Fortunately, in the Netherlands their number is limited.

DUO supplies Statistics Netherlands with annual data about individual pupils, by reference date 1 October. This reference date was chosen because school terms start at the end of August or the beginning of September. On 1 October late subscriptions are taken into account and the number of drop outs is minimal. In the feasibility study the following data were used:

- Citizen service number BSN
- Birth date
- Address

Data on Dutch diplomatic and military personnel outside the Netherlands

As pointed out in chapter 1, Dutch diplomatic and military personnel with a post abroad should deregister from the population register if the duration of stay outside the Netherlands exceeds two thirds of the forthcoming twelve months. The same applies to the partner and children travelling with them. Since it is not known if this is always done correctly, the Ministry of Foreign Affairs provided Statistics Netherlands with a file of all Dutch diplomatic personnel with a post abroad in 2012-2014.

Unfortunately, no permission was given for the provision of data about military personnel and persons employed in national service departments. Their number is limited, not exceeding 150 on an annual basis. The maximum term for military personnel on peace keeping missions, such as in Afghanistan or Mali, is six months. Both the usual residence population definition and the criteria for deregistration from the Dutch population consider these persons as a part of the Dutch population during these missions.

The following data were provided by the Ministry of Foreign Affairs:

- Citizen service number BSN (if registered in the population register by the time the data file was compiled)
- Birth date
- Date when the post in the foreign country started
- Date when the post in the foreign country ended

Missing data

Unfortunately, not all usual residents were covered by one or more registers above, such as:

- Illegal and undocumented persons;
- Household members of foreign workers without a paid job;
- Itinerants;
- Homeless people.

However, not all persons mentioned above are missing in the population register. To give an example, a homeless person may be registered at the address of a friend or a relative. Moreover, some groups may overlap: an illegal person may be homeless at the same time.

Other people who have left the country for over 12 months are not covered by any of these registers, such as:

- Inhabitants who have left the country to be employed abroad, and their household members;
- Students in higher education who pursue their studies abroad.

Obviously no accurate data are available about persons whose whereabouts are not covered by some register.

Variants and methods applied

Data from all data sources mentioned above were combined in one dataset. For each individual the start and end dates from each data source were combined such that the total presence time could be determined. Since there is no single combination of data sources that directly lead to the intended population figure, six variants were distinguished. The outcomes give an indication of the variation caused by the uncertainty regarding the underlying assumptions. The variants are as follows:

1. Both deregistration from the PR less than 365 days and in between registration in ER of less than 32 days were neglected. In case of incompatible data (example: an individual is registered in PR and receives monthly allowances abroad) the presence in the country was followed. People who reside in the country for more than 365 consecutive days are considered usual residents.
2. Both deregistration from the PR less than 365 days and in between registration in ER of less than 32 days were neglected. In case of incompatible data (example: an individual is registered in PR and receives monthly allowances abroad) the presence abroad was followed. People who reside in the country for more than 365 consecutive days are considered usual residents.
3. Both deregistration from the PR less than 30 days and in between registration in ER of less than 32 days were neglected. In case of incompatible data (example: an individual is registered in PR and receives monthly allowances abroad) the presence in the country was followed. People who reside in the country for more than 365 consecutive days are considered usual residents.
4. Both deregistration from the PR less than 30 days and in between registration in ER of less than 32 days were neglected. In case of incompatible data (example: an individual is registered in PR and receives monthly allowances abroad) the

presence abroad was followed. People who reside in the country for more than 365 consecutive days are considered usual residents.

5. People who reside in the country for more than 365 days (not necessarily consecutive) during two consecutive calendar years are considered usual residents. In case of incompatible data (example: an individual is registered in PR and receives monthly allowances abroad) the presence in the country was followed.
6. People who reside in the country for more than 365 days (not necessarily consecutive) during two consecutive calendar years are considered usual residents. In case of incompatible data (example: an individual is registered in PR and receives monthly allowances abroad) the presence abroad was followed.

Variants 1 and 2 comply with Regulation (EC) No 862/2007 on Community Statistics on Migration, variants 5 and 6 aim at adequately categorizing persons who regularly live in more than one residence during the year (in the country and abroad). The latter two do not require 365 consecutive days. Persons who meet the criteria of variants 1 or 2 automatically meet those of variants 5 and 6.

Variants 3 and 4 were introduced as a kind of contrast to variants 1 and 2, respectively. Compared to variants 1 and 2, the number of usual residents is decreased by some 17 thousand when the minimum migration period is 30 days (instead of 365). The number of short term emigrations is obviously bigger than the number of short term immigrants. Besides, migrants according to variants 1 and 2 are by definition implied by variants 3 and 4.

For some 4 thousand persons mutually incompatible data were found. In most cases these data refer to persons who are inhabitants of the Netherlands according to the population register data and who receive allowances at an address in a foreign country according to other data sources, which makes the quality of the PR data debatable. It turns out that about 1.5 thousand of them contribute to a decrease of the under coverage and another 2.5 thousand to an increase of the over coverage:

Usual residents, 1 January 2013, six variants

	Usual residents	Surplus with respect to population size	Under coverage	Over coverage
	x 1,000			
Variant 1	16,866.7	87.2	105.0	17.9
Variant 2	16,862.9	83.3	103.5	20.2
Variant 3	16,849.2	69.6	91.3	21.7
Variant 4	16,845.7	66.1	90.1	24.0
Variant 5	16,915.1	135.6	153.5	17.9
Variant 6	16,913.1	133.5	152.1	18.6

In all variants the size of the usual residence population turns out to be higher than the size of the population presented in the national demographic statistics, based on the population register. The biggest differences are shown by variants 5 and 6, referring to

persons who spent at least half of their time in the country, but not necessarily for 365 consecutive days. In addition to variants 1 and 2, there are some 50 thousand persons who spent a number of short term registered periods in the Netherlands where the intervals between these periods were considered to be too big, which led to the conclusion that one or more of these intervals were spent abroad, presumably in their home country. The majority are workers who have an on-and-off labour relationship with one or more employers. They leave the country when they are in between jobs.

Many Central Europeans do not succeed in obtaining long term contracts. Some have been in the country for many months or even years, but their contracts are not long enough to be entitled to be registered as residents in the population register. Thanks to the Employment Register they are identified as belonging to the usual residence population.

A limited number of persons who are registered in the population register are omitted from the usual residence population. This group mainly consists of immigrants who leave the country within twelve months after their arrival and persons who are registered as inhabitants but receive allowances abroad.

The most striking group that is not covered by these figures consist of undocumented immigrants. Clearly no register data are available. However, their numbers can be estimated by using capture-recapture methods, in particular the truncated Poisson regression model (e.g. Van der Heijden *et al.* (2015) estimated the number of undocumented immigrants in 2012 and 2013 at 35,530).

Variants 1 and 2 described above meet the requirements of Regulation (EU) No 1260/2013, variants 3-6 do not. The difference between variants 1 and 2 focusses on persons whose presence is contradicted by other sources. The choice between these two variants was based on the number of contacts between the register and the public. Allowances are usually paid once a month whereas contacts with the population register are usually less frequent. It is therefore argued that variant 2 meets the requirements of the usual residence population best. All persons whose presence in the Netherlands exceeded 365 days are included in this population, where persons whose presence is contradicted by other sources are left out. According to this variant, some 104 thousand usual residents were not covered by the population register, whereas some other 20 thousand were unjustly included. In the next section these figures are to be combined with the results from the capture-recapture method.

3.3 Combining the capture-recapture method and the micro register data method

The capture-recapture method and the micro register data method both have advantages and disadvantages. The capture-recapture method produces a reliable estimate of the under coverage, but this method is not fit for estimating the over coverage. The micro register data method produces estimates of both under and over coverage, but some data sources are lacking, so both phenomena are underestimated.

Firstly we focus on under coverage. According to the capture-recapture method the number of usual residents not included in the population register is in the interval [153.9, 185.9] (x 1,000). The centre of this interval (169.9 x 1,000) is considered to be the most plausible estimate. The micro register data method produces a much lower figure, 103.5 thousand. The difference between the two is explained by a number of factors. Firstly, some data sources were lacking. As was mentioned above, illegal and undocumented persons are the most striking group in this respect. Their number was estimated to be 35.5 thousand, which explains half of the gap. Among other groups for which data are lacking are partners of foreign workers who are not employed themselves. Their numbers are not known. The number of homeless persons not covered by other data is assumed to be low. A number of them are registered in the population register, for instance at the address of their parents or other relatives. Others are registered at the address of an organisation that provides medical or social assistance. For most homeless people who are not registered by the population register their presence in the country is not legal, which means that they are included in the 35.5 thousand mentioned above.

Improved economic conditions between January 2013 and March 2014 are another factor that add to some extent to the explanation why the capture-recapture-estimate is higher than the micro register data method estimate of the under coverage. The latter's reference date is 1 January 2013 and the capture-recapture method refers to 28 March 2014, almost 15 months later. In these months the number of workers (not all of them necessarily usual residents) from EU member states (excluding the Netherlands itself) and candidate member states increased by some 14.3 thousand. This increase is reflected by the public's confidence in the economic situation. In February 2013 it was at an all-time low, but it increased especially rapidly in Fall 2013. Household expenditures followed almost the same pattern.

Last, but not least, all figures relating to under and over coverage are based on a number of assumptions of which the validity can only be partially checked. The variants presented above provide plenty of proof for that. Moreover, as was pointed out before, the capture-recapture method results have quite some uncertainty attached to them. It follows that explaining the difference between the under coverage estimates can only be rough.

In order to arrive at the best estimate of the usual residence population the results of both methods were combined. The capture-recapture method is believed to produce the best under coverage estimate. Given the lack of some data sources, the over coverage estimate of the micro register data method was considered to be too low. Here it was assumed that the ratio of the two under coverage figures is also valid for the over coverage. Obviously it was not possible to check the validity of this assumption. The over coverage was therefore estimated as 33.2 thousand. So the surplus with respect to the 1 January 2013 population size based on the population register is estimated at 136.7 thousand, which is 0.8 per cent of the official population size. It follows that the usual residence population is estimated as 16,916.3 thousand on 1 January 2013.

3.4 Comparison of the 2010 and 2013 estimates

The figures on the Dutch usual residence population published by Eurostat so far (2014-2016) were based on an extrapolation of the 2010 estimate, assuming equal developments for registered and non-registered usual residents (Annex 5). There are two big differences with the estimates presented in this report:

- In the present estimates under coverage and over coverage were taken into consideration. The former estimates presented only under coverage.
- Since 2014, the capture-recapture method was improved in a number of steps. The methodology, the software and the estimation of the under coverage of the 0-14 years old were improved, thereby improving compliance with the assumptions underlying the CRC method. The figures presented in Gerritse et. al (2016) were based on a number of these improvements. In the current report the method was improved further. As a result the present width of the uncertainty interval surrounding the estimate is now much smaller than the 2010 one.

Today the original 2010 estimate is to be considered as rather rudimentary. Starting from 2017, the annual update of the usual residence population size will be based on the current method. Obviously, this will lead to a break in the series of figures and we propose to discuss with Eurostat how to deal with this.

The table gives both the extrapolated estimates based on 2010 data that are based on annually forwarded to Eurostat and the new estimates based on the methods described in the current report if a similar extrapolation procedure was applied.

Usual residents based on different sources and data

	Extrapolated CRC estimate based on 2010 data	Current estimate
	x 1,000	
1 October 2010	16,889.3	
1 January 2011		
1 January 2012		
1 January 2013		16,916.3
1 January 2014	17,082.7	16,966.4
1 January 2015	17,155.2	17,038.5
1 January 2016	17,234.7	17,117.5

The decreased interval width gives Statistics Netherlands more confidence in the new estimate than in the former one. However, given the uncertainties attached to the Crime Suspect Register used in the capture-recapture method, the missing data for some population groups and the assumptions that needed to be made when applying the CRC method and the micro register data method it must be stressed that neither the former estimate nor the new one can be considered to produce a definite reliable estimate of the usual residence population size. For this reason Statistics Netherlands

reserves the right to revise the estimates when new data or new data sources become available.

3.5 Regional data

Regulation (EU) No 1260/2013 not only demands the usual residence population size at the national level, but also on NUTS II and NUTS III level. However, application of the capture-recapture method only led to figures at the national level. This method could not produce data at the regional level. This is why data at the regional level were arrived at by application of the micro register data method, by retrieving the region were the residents were registered on 1 January 2013. Unfortunately, region was not available for all residents. There were about 91.7 thousand unknowns. Furthermore, the combination of the capture-recapture method and the micro register data method (variant 2) led to a surplus of another 169.9 – 103.5 = 66.4 thousand residents for whom the region of residence is not known, leading to a total number of 158.1 thousand unknowns. Obviously all unknowns are part of the under coverage (93%). See also Chapter 6 Datasets.

3.6 Vital events

The birth certificates issued by the Dutch local registrars indicate that no births are missing. This was not entirely the case with respect to deaths, where about 650 annual deaths in the Netherlands were found among people not registered in the population register. However, it is not known how many of these persons were usual residents. Generally speaking cause of death cannot give a conclusive answer about whether the deceased person was a usual resident or not, but in some cases some clue is given. Indeed, it may be argued that persons who died in homes for the elderly or nursery homes might be residents of the Netherlands long enough to be considered as usual residents. Their annual total does not exceed 40. On the other hand, the costs of such homes are paid by the government. It follows that it is almost impossible to reside in such a home without being registered in PR, which contradicts the argument.

4. Evaluation of the feasibility

Key elements

Three key elements of the usual residence population concept are mentioned in Regulation (EU) 1260/2013 on European Demographic Statistics, viz. national reference date, main rule(s) applied and duration of stay and modality of duration. The feasibility study points out that Statistics Netherlands is able to produce population figures for any reference date, including 31 December and 1 January. Statistics Netherlands clearly satisfies this key element.

The same applies to the second key element. The population register, which is the main source for producing demographic figures, does not satisfy the 12 months criterion. This is why two methods were developed to satisfy this criterion, as pointed out in the previous chapter. Since no sources are available that lead directly and unambiguously to the usual residence population, a number of assumptions were needed in both methods. In order to have an indication of their validity, a number of variants were developed and confidence margins were estimated.

Application of the 12 month criterion may refer to intended or realised stay. Statistics Netherlands cannot produce intended stay figures on a 12 months basis. However, because the number of usual residents was established for the reference date 1 January 2013, it was possible to produce duration on the basis of realised stay . The two full calendar years 2012 and 2013 could be taken into consideration to measure the actual number of days spent in the country. An interesting question which cannot be answered is whether population according to intended stay of at least 12 months would have been closer to intended stay as included in the PR or to realised stay of minimal 12 months.

The high quality of the registers with respect to the people who are officially allowed to reside, work, study , ... in the Netherlands makes this approach useful and fruitful. On the other hand it is acknowledged that these registers do not cover persons whose presence is not officially allowed. The latter leads directly to the third key element. It is impossible to have corrections for each of the so called 'problematic' groups at the individual level. For obvious reasons there are no accurate, complete and reliable register data available for a number of groups. Instead, the capture-recapture method aims at estimating the under coverage for the whole of these groups. Unfortunately, this method does not permit to estimate over coverage as well.

Accuracy

We hoped that the results by the capture-recapture method and the micro register data method were more or less the same, creating more confidence in the results. The methods are completely different, and both have advantages and disadvantages. The capture-recapture method enables us to obtain data of the total under coverage of usual residents at the national level, no matter what the cause is of not being covered by the national demographic statistics. However, the method is not fit for estimating over coverage. The micro register data method suffers from the fact that no appropriate data sets are available for all groups. It follows that this method underestimates both under and over coverage.

The possibilities to estimate over coverage are limited. As said, the capture-recapture method does not allow such estimates since only Dutch data sources are available. Although a correction is made for immigrants whose stay in the country is less than twelve months, in fact the micro register data method has to deal with the same problem. Here foreign sources with data about Dutch citizens who work or pursue their studies abroad would be useful. Lack of access to foreign data sources, if any, inevitably leads to underestimation of the over coverage. From this point of view the usual residence population may be believed to be too high. On the other hand, the micro register data method underestimated under coverage as well. For example, no data sources on family members of foreign workers are available. It is not known which of these two effects is the biggest. In this feasibility study it was assumed that the underestimation in terms of percentage of under and over coverage are equal.

In the micro register data methods, an indication of the stability of the data can be achieved by slightly altering the assumptions. If, for example, short term migrants were taken into account, the number of usual residents fell by some 17 thousand. By contrast, if the requirement of 365 consecutive days is dropped, under coverage increased by almost 50 thousand. The effect of these assumptions on over coverage is much smaller, some 6 thousand at the most.

Costs and feasibility / Future procedure

Carrying out both the capture-recapture and the micro register data method turned out to be costly and very time consuming. Obtaining data from the Ministry of Foreign Affairs and the Immigration and Naturalization Service took six and twelve months, respectively. The data analysis of these sources was relatively simple and straightforward. By contrast, some data used in the capture recapture method were less stable than expected, demanding much more time and commitment than planned. Given the different outcomes of the two methods, combining the results and analysing the differences also took time. The conclusion is that at present it is impossible to carry out this procedure on an annual basis. For validation reasons it should, maybe partially, be repeated after a number of years. During the years in between, a less time-consuming and more simple method should be implemented. A cost-benefit analysis will be carried out in order to establish the most appropriate approach for the years to come.

The first step taken then is analysing the fragmented data sources, in order to have an indication of the assumed stability. It must be kept in mind that all data sources were designed and kept for non-statistical purposes. If data sources are no longer available, or when major changes have taken place, the validity of the estimation method must be re-investigated.

The acquisition of data sources took up to a year. The main reason was that the registration holders had to develop a strategy to arrive at the data asked for by Statistics Netherlands. Furthermore, those organisations had to plan these one-off activities beyond to the tasks they already had. This turned out to be time consuming, and in some cases low priorities were set. Fortunately, Statistics Netherlands is entitled by law to get access to every governmental data source. If necessary, structural arrangements can be made with data owners aiming at short handling times.

5. Timetable

In the former section it was argued that applying the capture-recapture method and the micro register data method and combining the results appears to be very time consuming. At present the main conclusion is that it is impossible to carry out this procedure on an annual basis. A cost-benefit analysis will be carried out in order to establish the most appropriate approach for the future. Here decisions have to be made about which elements of both methods have to be repeated after how many years, how to produce the required figures in the years where the extended methods have not been used and whether revision procedures are possible and needed for older years.

If for the sake of utter simplicity, and following the procedure used so far (see Annex 5), it is assumed that the 2013 ratio of the official population size derived from the population register and the estimated usual residence population is also valid for 2014-2016, the usual residence population size for these years is as follows:

Usual residence population, 1 January 2013-2016

	Population size under national concept	Usual residence population
	x1,000	
2013	16,779.6	16,916.3
2014	16,829.3	16,966.4
2015	16,900.7	17,038.5
2016	16,979.1	17,117.5

It is acknowledged that the usual residence population size forwarded to Eurostat so far (see Annex 5) needs revision. Obviously, application of the new method will lead to a break in the series of figures. We propose to discuss with Eurostat how to deal with this.

In spite of the increased confidence in the new estimate, as pointed out in section 3.4, it must be stressed that neither the former estimate nor the new one can be considered as very reliable. For this reason Statistics Netherlands reserves the right to revise the estimates when new methods and techniques or new data sources become available.

6. Datasets

Following Article 3 of Regulation (EU) No 1260/2013 here data are to be presented on the usual residence population by age, sex and region of residence. Application of both the capture-recapture method and the micro register data method led to known figures for over 99.0 per cent of that population. The figures were arrived at in two steps, as follows. The micro register data method led to some 8 thousand persons of whom age and sex could not be established. For region of residence the number of unknowns is much bigger: 91.7 thousand.

Usual residents according to variant 2, by age, sex and region of residence, 1 January 2013

	Total	of which		
		available	not available	
	x 1,000			%
Age	16,862.7	16,854.9	7.9	0.0
Sex	16,862.7	16,854.7	8.0	0.0
NUTS level 1	16,862.7	16,862.7	-	-
NUTS level 2	16,862.7	16,771.1	91.7	0.5
NUTS level 3	16,862.7	16,771.1	91.7	0.5

The total number of usual residents was determined by combining the results of the micro register data method and the capture-recapture method, leading to 16,916.3 thousand. The surplus with respect to the variant 2 figure by age, sex and region of residence (66.4 thousand) is not known either. This leads to the following:

Usual residents by age, seks and region of residence, 1 January 2013

	Total	of which		
		available	not available	
	x 1,000			%
Age	16,916.3	16,842.0	74.3	0.4
Seks	16,916.3	16,841.9	74.4	0.4
NUTS level 1	16,916.3	16,916.3		
NUTS level 2	16,916.3	16,758.3	158.1	0.9
NUTS level 3	16,916.3	16,758.3	158.1	0.9

The table shows that for 4 in 1,000 usual residents age and sex are not known. The number of unknowns with respect to region of residence is 9 in 1,000.

In order to avoid unknowns, imputations should be needed. The introduction of such kind of imputations and the demands for comparability over time should be taken into account in the cost-benefit analysis needed to make decisions about future practice.

There is hardly no information at all available which can be used to produce the specific numbers of vital events among the usual residence population other than already available in the population register in the Netherlands.

References

Bakker, Bart, and Kees Prins (2016), Two methods for estimating under coverage in administrative sources. Paper presented at the Conference of European Statistics Stakeholders (CESS) meeting, 20 October 2016, Budapest, Hungary.

Bishop, Y.M.M., S.E. Fienberg and P.W. Holland (1975), *Discrete multivariate analysis*, (Cambridge, MA: MIT press).

Di Consiglio L. and Tuoto T. (2015) Coverage Evaluation on Probabilistically Linked Data, *Journal of Official Statistics*, Vol. 31, No. 3, 2015, pp. 415–429.

Fienberg, S.E. (1972), The multiple recapture census for closed populations and incomplete 2k contingency tables, *Biometrika*, 59, pp. 409-439.

Gerritse, S.C., B.F.M. Bakker, P.-P. De Wolf and P.G.M. Van der Heijden (2016), *Under coverage of the population register in the Netherlands, 2010*. Discussion paper 2016-02 (The Hague / Heerlen: Statistics Netherlands).

Herzog, T. N., F.J. Scheuren, W.E. Winkler (2007). "Data Quality and Record Linkage Techniques", Springer, New York.

Van der Heijden, P.G.M., J. Whittaker, J., M. Cruyff, B. Bakker and H. Van der Vliet (2012), People born in the Middle East but residing in the Netherlands: invariant population size estimates and the role of active and passive covariates, *Annals of Applied Statistics,* 6, pp. 831-852.

Van der Heijden, P.G.M., M. Cruyff, G.H.C. van Gils (2015*), Schattingen illegaal in Nederland verblijvende vreemdelingen ('Estimating the number of non-Dutch nationals whose stay in the Netherlands is against the law') 2012 – 2013* (The Hague: WODC Ministry of Security and Justice). Only in Dutch.

Glossary

Administrative entrance	Entrance in PR after administrative removal
Administrative removal	Removal from PR when whereabouts are unknown
BSN	Dutch unique personal identification number
CRC	Capture-recapture estimation method
CSR	Crime Suspects Register
DUO	Dutch Education Service
ER	Employment Register
Over coverage	All usual residents who are not registered in PR
PR	Population Register
PROBAS	Register on staff members of international organisations
Under coverage	All persons in PR who are no usual resident

Annex 1 Template for describing the national population definition used in the demographic statistics_NL

National population definition (as self-reported by country)

Usually resident population	Present/de facto population	Registered population	National population	Legal/de jure population
N	N	Y	N	N

National reference date of the population counts/estimates

	Y/N
End of the year	N
Beginning of the year	Y
Mid-year	N
Other	N

Main criteria applied for national population counts/estimates

	Y/N
Presence in the territory of the country at a specific date	N
Having lived in the territory of the country for a given period of time (actual stay)	N
Intention to stay in the territory of the country for a given period of time	N
Legal right of stay (including citizenship)	N
Listing in national register(s)	Y
Other (please specify)	N

Duration of stay and modality for the national population counts/estimates

Modality/Duration	***3 months***	***4 months***	***6 months***	***12 months***	***at least 12 months***
Continuous					
Most of the time					
Minimum period		x			
Not defined					
Other					

Deregistration of emigrants

	Y/N	If yes, please specify
Legal obligation to deregister when leaving the country	Y	A person has to deregister if it is expected that he/she will stay abroad for at least two thirds of the forthcoming twelve months. This requirement is the only incentive to deregister.
Incentive to deregister when leaving the country	N	

Inclusion/exclusion of selected groups for population data at the NATIONAL level

Population groups		Included in population data at national level
A	Persons regularly living in more than one residence during the year (in the country AND abroad)	Yes, provided their stay outside the Netherlands does not exceed eight in twelve months.
B	Persons (civilian residents) working abroad (crossing a frontier daily or weekly to work in another country but returning home every day or at weekends)	Yes, provided their stay outside the Netherlands does not exceed four in seven nights.
C	Primary and secondary school pupils and students away from home during the school term and pursuing their education abroad	Yes, provided the net school term does not exceed eight in twelve months.
D	Tertiary students away from home during the school term and pursuing their education abroad	Yes, provided the net school term does not exceed eight in twelve months.
E	Persons living in an institution abroad	No.
F	National military, naval and diplomatic personnel and their families, located outside the country	Yes, provided their stay outside the Netherlands does not exceed eight in twelve months. This is usually the case for military personnel.
G	Foreign military, naval and diplomatic personnel and their families, located in the country	No, but they are allowed to be registered and included in population data. About one in three does so.
H	Foreign persons working for international organizations (not including diplomats or military forces) located in the country	No, but they are allowed to be registered and included in population data. About one in three does so.
I	Nationals working abroad for international organizations	Yes, provided their stay outside the Netherlands does not exceed eight in twelve months.
J	Nationals born abroad in the last 12 months	Yes, provided their stay outside the Netherlands does not exceed eight in twelve months.
K	Children who alternate two places of residence (one abroad and one within the country)	Yes, provided their stay outside the Netherlands does not exceed eight in twelve months.
L	Persons whose stay in the country is exactly one year	Yes.
m	Persons whose stay in the country is less than one year, even if for a single day (short-term migrants)	No, unless their stay in the Netherlands exceeds two thirds of the forthcoming six months.
N	Holders of temporary residence permits (and their families) staying in the country for more than 12 months	Yes.
O	Persons who may be illegal, irregular or undocumented migrants	No, but there are some exceptions.

P	Asylum seekers	Yes, provided they have been in the country for six months or over.
Q	Persons who have applied for or have been granted refugee status or similar types of international protection	Yes, provided they have been in the country for six months or over.
R	Nomads or other travelling people	Yes, provided their stay outside the Netherlands does not exceed eight in twelve months.
S	Homeless or roofless people	Yes, provided their stay outside the Netherlands does not exceed eight in twelve months.
T	Persons living in remote areas	Yes.

Inclusion/exclusion of selected groups for population data at the REGIONAL level

Population groups		Criteria used to decide on the region for the place of residence
a	Persons regularly living in more than one residence during the year (in different regions)	The region where the persons spend their daily rest most of the time.
b	Persons (civilians) working in another region of the country but returning home every day or on the weekends	Family home.
c	Primary and secondary school pupils and students away from home during the school term and pursuing their education in another region	Family home.
d	Tertiary-level students away from home during the school term and pursuing their education in another region of the country	Family home.
e	Persons living in an institution in another region of the country	Institution address. But: persons living in institutions have the right to choose a different address, see Annex 2.
k	Children who alternate between two places of residence (in different regions)	The region where the persons spend their daily rest most of the time.
r	Nomads or other travelling persons	The region where the persons spend their daily rest most of the time. Usually their address is with relatives, friends or institutions.
s	Homeless people	The region where the persons spend their daily rest most of the time. Usually their address is with relatives, friends or institutions.

Annex 2 Population register; definitions and practices

The population registers in the Netherlands are set up and maintained by the Dutch municipalities[5] according to the national Basic Population Register Act (PR Act). This gives strict rules for keeping the population registers. This guarantees that the municipalities apply the rules in a uniform way. Each registered person is given a unique personal identification number (pin), known as Citizen Service Number (in Dutch: BSN). It guarantees that each resident is registered in no more than one municipality. BSNs are strictly individual and are never attributed to anyone else, even after someone has died or emigrated. What is more, no data are erased from the population register after someone had died or emigrated. In spite of the fact that data are kept in perpetuity, we use the term 'deregistration' in this report. This is to indicate that someone is no longer a resident in the sense of the population register.

Persons need a BSN to get access to various services in the Netherlands. A number of Dutch agencies and national organizations use BSN for their specific tasks when registering the personal data of privileged persons. Important examples are:

- National tax authorities: BSN is mandatory when paying taxes, and for receiving compensations for high rent and health insurance premiums. These allowances are attributed to people whose income is considered too low. For privileged persons, including foreigners working on Dutch soil as diplomatic, consular or military officials, or in an international organisation, it is mandatory for claiming their fiscal privileges.
- Social security service: BSN is mandatory for receiving child allowance.
- Health care: BSN is mandatory for the administration of hospitals, dentists, medical doctors, medical insurance companies etc. Also the BSN is the single point of reference in the communication in the health care sector.
- Education: It is mandatory for Dutch schools, day care facilities and after school day care to use BSN in their administrations.

These examples indicate that it is nearly impossible to live a regular life in the Netherlands for people who are not registered. This is why there are great incentives for being registered.

BSN serves as the most important linkage key in governmental and many non-governmental registers, facilitating linkage of registers from a wide range of domains, such as taxes, labour, social security, education and health care. What is more, the linkage quality is extremely high since the uniqueness of BSN is guaranteed.

Once a BSN is assigned, the registered person keeps this pin for the rest of his or her life and beyond. The linkage between individual and BSN is never broken. The only exception is when a BSN is accidently assigned to two different individuals and when an individual accidently is assigned two different BSNs. Fortunately, this very rarely happens.

[5] It is emphasized that the population registers are not kept by Statistics Netherlands. From these registers Statistics Netherlands obtain data that are needed for demographic and other statistics of the population of the Netherlands.

Registration and deregistration; rules

The PR Act gives rules for registration and deregistration. Everyone who enters the Netherlands is registered as an resident (immigrant) provided

- His/her stay is legal according to the Immigration Act on people who do not have the Dutch nationality;
- the expected stay is at least two thirds of the forthcoming six months;
- the person is properly identified. The latter means that a valid passport or other official paper is shown for identification.

Every child born in the Netherlands whose mother is registered as a resident is also registered as a resident. Children who are born abroad to a mother who herself is registered as a resident of the Netherlands are registered, provided the children will live in the country. The child is entered in the population register of the municipality where the child lives. Usually this is the municipality where the mother is registered as an resident. Still-births are not registered in the population register.

Deregistration as an resident takes place when someone dies, or leaves the country and intends to stay abroad for at least two thirds of the forthcoming twelve months (emigration). The rules concerning registration and deregistration make no distinction between Dutch and non-Dutch nationals. Of course, the Immigration Act is applicable to non-Dutch nationals only.

At first glance the expression ‘two thirds of the forthcoming six or twelve months’ may seem a little odd. In practice the period usually coincides with the forthcoming four or eight months. The reason why the expression was adopted in the PR Act is to have it applicable to people who frequently cross the border for a short stay, either abroad or within the country.

There are two exceptions to the rules for registration. The first applies to so-called privileged persons mentioned above. Since they enjoy a special 'privileged' status and are not considered foreigners under the auspices of the Immigration Act, they are given the choice whether or not to be entered in the population register. Asylum seekers form another exception. Their registration takes place only six months after their arrival in the Netherlands, irrespective of their expected duration of stay, unless they are granted a residence permit within six months. In that case they are registered when the residence permit is granted. However, children born to asylum seekers who are not yet registered are registered directly after birth. This leads to the somewhat odd situation that a new-born baby is registered whereas the parents and siblings are not (yet).

Asylum seekers are allowed to stay in the country as long as no formal decision about their resident status is taken. Those who are ordered to leave the country within six months after their arrival are not registered as residents. Others are registered six months after their arrival in the country. If they have to leave the country, deregistration takes place. If they do not leave the country prolonged residence is illegal.

The rules with respect to registration and deregistration are clearly not symmetrical. To be registered, the intended duration of stay for entries is much shorter than for departures. To give an example, a resident of the Netherlands who spends almost eight

months of the year in Spain remains registered as a Dutch resident, a resident of Spain who spends at least four months in the Netherlands is registered here as an resident.

Registration and deregistration are never adjusted for the realized duration of stay. If, for example, a person who intends to stay in the country for six months, say, is registered, the registration is not annulled if he or she leaves the country after six weeks. The same applies to emigration. If a person returns to the Netherlands within eight months after deregistration, the registered emigration is not annulled. If, on the other hand, someone is not registered because the expected stay is too short, and the actual stay exceeds four months, that person's entry is not retrospectively registered. What is more, if the prolonged stay is not likely to exceed two thirds of the forthcoming six months, that person is still not registered. The latter may occur in practice, for instance, if someone has a series of short term labour contracts.

Registration and deregistration; examples

Below some additional information is given with respect to registration and deregistration with respect to the groups mentioned in Annex 1.

People who work abroad but return home every day are registered as residents of the Netherlands. If, however, workers spend five, six or seven nights a week abroad, deregistration should take place. So people who work abroad and return only during the weekends (e.g. only Friday and Saturday night) should no longer be registered as residents.

A similar reasoning holds for people who work in the Netherlands and regularly return to their homes abroad. If they spend at least three nights a week in the Netherlands, they must be registered as residents of the Netherlands (provided their stay is legal according to the Immigration Act). If they spend five nights a week or more at their home abroad, they are not registered as residents of the Netherlands.

Determining the exact number of nights spent or to be spend in the Netherlands is clearly not always possible. This applies to the workers and to the municipal authorities. People may not be able to predict the exact number of nights spent at home or abroad. But even if workers know that they will only spend one night a week in the country, they may prefer to be registered as a resident with the rest of the family and will not always provide the municipal authorities with the proper information that is needed to keep the registers up-to-date.

The same reasoning holds for primary and secondary school pupils and students away from home during the school term and pursuing their education abroad. Similar to the workers' example, parents may prefer to keep their children registered as a part of the family and will do their best not to exceed the duration of stay associated with emigration. In many cases this is plausible. If, roughly speaking, the pupil spends the holidays and half of the weekends at home, the number of days that corresponds to two thirds of twelve months is not exceeded.

Foreign pupils who spend their school term in the Netherlands usually spend enough nights in the Netherlands to be registered as residents. If their parents fail to inform the municipal authorities, those authorities will usually receive the information from others, such as the school governors. In those cases the parents can be forced to have their child registered as an resident in PR.

Foreign students at a university in the Netherlands are registered as residents upon their arrival in the country. Most universities in the Netherlands cooperate with the municipalities to facilitate proper registration. To be precise, on registration days registration is done at the university instead of the town hall. Foreign students who are new in the Netherlands are encouraged to participate in such events. In this way being registered is made easy, reducing the probability of students not being registered. Unless the duration of stay is less than four months, registration is obligatory.

During registration in the PR, a foreign student is asked the expected date of departure. Unless the student notifies the municipal authorities of an early departure, a reminder is sent when the planned emigration date comes closer. In this manner the authorities aim to keep the number of incorrectly registered persons as low as possible. Of course, if circumstances change the student is usually allowed to stay in the country longer.

Dutch residents who study at a university abroad are deregistered if the expected duration of stay in the foreign country exceeds two thirds of the forthcoming twelve months. This is usually the case, even if the holidays are spent with the family in the Netherlands.

National diplomatic personnel and their families, on location outside the country, are usually deregistered. However, this usually does not apply to national military and naval personnel, given the fact that the maximum duration of stay of military personnel in peacekeeping operations is six months, which is very rarely exceeded.

Persons whose stay is not legal, are not registered as residents. An illegal stay can arise at various stages. A person's stay may be illegal at the arrival in the country, for instance if lacking the required visa. People who do not need a visa are allowed to remain legally in the country for 90 days, so they stay illegally after the 91st day. Illegality also arises when people do not leave the country when their visa or temporary residence permit has expired. They may have been registered as residents when their stay was legal, but as soon as it becomes clear to the municipal authorities that their stay is no longer legal they are deregistered.

Homeless people who stay in the country legally are usually registered. Their registered address may depend on the factors that led to their being homeless, such as their health history. Homelessness may be caused by mental illness, addiction to alcohol or drugs, the end of a relationship, bankruptcy or a combination of causes. Some are registered at the address of an health institution, others at the address of a relative or friend, or at an address of an organization for the homeless. When they travel from town to town, they obviously may be registered in a region where they rarely or never are.

The number of itinerants is very low in the Netherlands. There is usually no distinction with homeless people. People who do not have a regular place to stay are registered in the municipality where they usually remain. People who are constantly on the move, for whom there is no municipality where they usually remain, are registered by the municipality of The Hague.

Holders of temporary residence permits (and their families) are registered as residents.

Place of residence

For each registered person, the place of residence (address) is registered. According to the PR Act, this is the place where someone lives. If someone has more than one address in the country, registration applies to the address where the persons resides most of the time. As to addresses abroad the rules for registration apply. In fact, when people have an address in the Netherlands and an address abroad, they are registered as residents if they spent at least one third of the time in the Netherlands.

People who have more than one address in the country, are registered at the address where they reside most of the time. For example, someone who spends four days / nights a week at one address and three days at the other address, is registered at the former address.

People who work in another region of the country and return home every day, are registered at the home address. If they return only at the weekends they obviously spend most of the time in the working region, so they should be registered there. However, similar to workers employed abroad, these people may not always provide the correct information to the municipal authorities. Many will prefer to be registered at their family's home address. But in some cases being registered at separate addresses is beneficial, given the rules on rent and health care subsidies. Indeed, the total income of the persons registered at an address is crucial for those subsidies to be granted.

Primary and secondary school pupils and students living away from home during the school term and pursuing their education in another region of the country, are usually registered at the home address. When they spend most of the time in the region where the school is, they should be registered there. Again, parents will not always provide the correct information to the municipal authorities. The municipal authorities may decide to contact the school governors in order to get the pupils and students properly registered.

Students in higher education are registered at the address where most of the time is spent. In order to facilitate correct registration many municipalities and universities organise registration days when the registration sessions are held at the university instead of the town hall.

People living in a health care, child care or penal institution are not obliged to be registered at the institution's address. Directors of these institutions must inform them of their right to be registered elsewhere. In fact, they are allowed to choose an entirely different address during their stay in the institution as long as the person who actually lives there is prepared to pass on their mail.

Public's support is compulsory

In order to help the municipalities in keeping the population registers up to date, the population is obliged to keep the municipal authorities informed about births, deaths and changes of address, including immigration and emigration. If people fail to do so, the municipal authorities are obliged to take formal action in order as yet to obtain the required information. Not informing the authorities, or deliberately providing false information, can be penalized with a fine up to €325.

Despite these obligations and penalties, in some cases the registered data do not match with reality. Such differences are usually the result of persons not informing the municipal authorities of their arrival, departure or move elsewhere within the country. Immigrants and emigrants have to inform the municipal authorities of their arrival or departure, but in practice a number fail to do so, or do not do so on time. This may be due to negligence or obstruction, i.e. persons forget to report their arrival or departure, or are reluctant to do so. But certainly not all non-reporting is due to human failure. For example, workers who come to the Netherlands for three months, say, are not obliged to inform the municipal authorities, since registration takes place only if they expect to stay for more than four months. If they find another job for another three months after the first job finishes, informing the municipal authorities is still not necessary. The reason is that the law does not permit retroactive registration. In this manner a person may live in the Netherlands for quite some time without being registered. The impossibility of retroactive registration also may lead to belated registration, for instance when people intend to stay for a short period time and then decide to prolong their stay.

Objections against being registered may be related to avoiding contacts with tax authorities or other officials. Sometimes employers or landlords forbid people to be registered by threatening to dismiss them or lose their home. Of course these things cannot be tolerated, and when discovered measures are taken to have people registered as yet. It must be stressed that not being registered is not equal to illegal stay. For instance, nationals from other EU member states who stay in the Netherlands and who do not register is entitled to be in the country. Their labour may be illegal if they have an unofficial job, but their presence is not.

Some delayed or non-registrations are due to negligence. People who move to the Netherlands usually have a lot on their minds and, not surprisingly, not all of them are aware of the obligation to register upon arrival. Others are not used to being registered in their country of origin which may explain their hesitation to do so in this country. Municipal authorities must take formal action to have these people registered. If obstruction is proven, the person faces a maximum fine of € 325. In practice the number of fines is very low.

Unnotified emigration

About one in three persons who leave the country do not notify the municipal authorities of their departure. When the authorities find out that someone is 'missing', the PR Act stipulates that they must investigate his or her whereabouts. Doubts about a person's whereabouts may arise when another person registers at the person's address and finds the place unoccupied, when taxes are not paid and the fiscal authorities are unable to find the individual, etcetera. The municipal authorities are then obliged to investigate the person's whereabouts. As a first step the municipal authorities send a letter to the registered address urging the addressee to inform them of the new address. If a new address is reported by other authorities, such as the social security office or the tax department, the municipal authorities may send a letter to that address as well. If the addressee does not respond within three weeks, the municipal authorities must take further actions by enquiring with the neighbours, the employer, relatives, the landlord etcetera. A drawback of this procedure is that these persons have the right not to respond. On the other hand, government agencies are obliged to share the information they have. Another possibility that can be combined with these actions is sending municipal officials to the registered address in order to

find out who is actually living there. Social media are another means to collect data about the whereabouts of the missing person. It goes without saying that municipal authorities are urged to carry out such investigations as thoroughly, yet as fast as possible.

If the investigations lead to the missing person's new address, it is registered in the population register. The best thing is to contact the individual so he or she can confirm this new address, but if this is not possible the authorities are forced to take formal action and register the new address without the person's approval. If, however, no information about the person's whereabouts can be obtained, the assumption is that the person must have left the country. However, before arriving at that conclusion, it must be clear that the investigations were carried out thoroughly during a four week period. What is more, the missing person must be officially informed that deregistration is planned. This can be done by sending a letter to an assumed new address. Or by publishing an announcement in the local newspaper or in other media, enabling the public to provide information about the person's whereabouts. After another fruitless four weeks deregistration takes place, some eleven weeks after the first letter to the last address. The new registered address is qualified as 'unknown' in a country qualified as 'unknown'. Sometimes this is referred to as 'departure to an place unknown to the authorities'.

The foregoing makes clear that a person whose whereabouts are not known, is given ample time and opportunities to respond to the actions taken by the municipal authorities. At the same time there is a drawback, which has to do with the PR Act decree that an address in the country must be registered for every resident of the Netherlands. The law does not allow people to register at a fake address of be without address. This means that until deregistration finally takes place, the person continues to be registered at the address even when it is completely clear that the person does not live there anymore. Only when the municipality decides at last that the whereabouts are definitely not known, the missing person is administratively emigrated to a country 'unknown'. Although the municipal authorities try to carry out the investigations as fast as possible, the minimum period of investigation is about eleven weeks, so during these weeks the registered address usually differs from the person's actual whereabouts. Also this person stays registered as an resident even if that person has actually moved abroad.

Annex 3 Description of the linkage process

For the capture-recapture process three registers have to be linked. These are the Population Register (PR), Employee Register (ER) and the Crime Suspects Register (CSR). Linkage is performed in two steps. First, the PR and ER are linked. This results in a new data set containing records (both linked and not-linked) from both data sets: the PR/ER. This new data set is then linked to the CSR.
Capture-recapture assumes perfect linkage. Therefore, linkage was performed in two steps. First the registers are linked deterministically using a unique person's ID present in all registers. The remaining records are then linked probabilistically (see Herzog *et al.*, 2007) using keys such are date of birth, gender and address information (postal code, street, etc.). Selections on place of residence and nationality were performed after the linkage process: to derive these, information from all three registers was combined.

Linkage of Population Register and Employee Register; description of the sources and data preparation

Population Register
The PR dataset used contains all persons registered in 2014. It also contains all changes registered in this period and before including address changes. In order to reduce the amount of data, we selected only records of persons aged between 15 and 65. All address changes between 1 June 2013 and 31 December 2014 were included. This resulted in 17,366,596 records.

Employment Register
The ER dataset contains information on all persons with a job in 2013 and 2014. All known addresses between 1 June 2013 and 31 December 2014 are included. Addresses belonging to persons whose unique person's ID could not be found in the PR were cleaned up, because these are addresses as registered by the employer and often contain errors.

Clean-up of addresses
The address fields consist of the following fields: street name and house number including addition, postal code (the Dutch postal code consists of 4 digits followed by 2 letters), and city. In principle the postal code plus house number uniquely identifies an address. However, not all records contain a valid postal code and some records contain errors in the postal code. Therefore, the addresses are first cleaned up using the Dutch Building Register (BAG) that contains all valid addresses.
In order to clean up the addresses the street names are first standardised:

- Additions to the house numbers (e.g. '-a', '/A', '-II') are standardised (e.g. 'a', 'a', '2')
- Non alpha-numeric characters are removed and street name is converted to lower case.
- Common terms such as 'burgemeester' ('mayor') and 'koningin' ('queen') are removed .
- Short (1-2 character) words are removed.

- The endings of street names are standardised (e.g. convert ‘straat’, ‘str’, ‘strjitte’, ‘strt’, and ‘st’ to ‘S’).

Addresses are matched using probabilistic linkage to addresses in the BAG and street name, postal code and city are taken from the BAG.

Record linkage

In the first step records from the ER are linked to the PR using the unique identifier. The remaining records from the ER are linked, using probabilistic linkage, to the complete PR and not just to that part of the PR that is not yet linked to the ER. This is because people can have more than one job. Therefore, an unlinked record from the ER can still belong to a record from the PR that is also linked to another record in the ER.

The following variables were used in the probabilistic linkage: date-of-birth, gender, postal code and house number. Blocking was performed on postal code and date-of-birth separately. The remaining variables were used as linkage keys. Using two blocking variables increases the possibility that a record with an error in one of the blocking variables can still be linked. This results in two linked data sets, one for each blocking variable. The two data sets are combined and de-duplicated: a record from the ER can only belong to one person from the PR. The pair with the highest linkage probability is selected. Linkage probabilities were determined using the standard EM-algorithm (Herzog *et al.*, 2007). The results from the linkage are presented in tables 0.1 and 0.2. Most records from the ER could be linked to the PR. Only a negligible number was linked using probabilistic linkage. Finally the two registers are combined into one register (PR/ER).

Table 0.1. Overview of the results from the linkage of the Population Register (PR) with the Employee Register (ER)

	PR	ER
	x 1,000	x 1,000
Number of records	17,368.0	7,833.0
Number linked	7,157.0	7,648.0
	%	%
Percentage linked	41.2	97.6

Table 0.2. Overview of final combined register PR/ER by type of linkage

	x 1,000
Number of records in combined register	17,546.0
of which unlinked PR	10,210.0
of which unlinked ER	179.0
of which Linked using deterministic linkage	7,156.0
of which linked using probabilistic linkage	0.9

Linkage of Crime Suspect Register; description of the Crime Suspects Register and data preparation

As discussed in the previous section, the records from the PR and the ER are combined into one register PR/ER. The new register is then linked to the Crime Suspects Register (CSR). The CSR contains crime suspects that are registered as such by the police. The CSR consists of two parts. One part had been linked by the police to the PR. For these records only the registration number used in the PR, which can be translated to the unique identifier, is known. For the remaining records personal information such as date-of-birth, gender and address is available which can be used in the probabilistic linkage process. The address information is cleaned using the same procedure as used for the ER.

Record linkage

All records from the CSR in the part that was not linked to the PR by the police were used in a probabilistic linkage between the CSR and the combination PR/ER. The CSR records were linked to the complete PR/ER. The linkage keys in the CSR contain a large number of errors (see '*Data Quality Issues*' below). Because of the size of the registers it is necessary to work with blocking variables in the linkage process. However, records that differ on their blocking variables will never be linked. In order to reduce the risk of this, different combinations of blocking variables and linkage variables were used, as summarised below:

1. Using postal code as the blocking variable and the following variables as linkage variables: year of birth, month of birth, day of birth, gender and house number.
2. Using date of birth as blocking variable and the following variables as linkage variables: gender, postal code, street name, city and house number.
3. Using the combination city, year of birth and month of birth as blocking variable and the following variables as linkage variables: gender, street name, house number, day of birth.
4. Using the combination city and year of birth as blocking variable and the following variables as linkage variables: gender, street name, house number, month of birth and day of birth. Blocks with more than 50 records in the CSR are excluded as these would give too many possible pairs.

5. Using city as blocking variable and the following variables as linkage variables: gender, street name, house number, year of birth, month of birth and day of birth. Blocks with more than 80 records in the CSR are excluded as these would give too many possible pairs.

Each of these five linkage steps results in a set of linked record pairs. These pairs are combined into one data set after which pairs are selected. Records from the CSR can only be linked to one person in the PR/ER. Persons from the PR/ER register can be linked to multiple records from the CSR because the CSR can contain multiple records belonging to the same person. The selection was performed in such a way that the total posterior m-probability (which gives the probability that a pair is a true match) is maximised.

Results

The results of the record linkage are presented in table 0.3. Of the 482 thousand records in the CSR that needed to be linked probabilistically to the PR/ER eventually 73 thousand records could be linked. The remaining 409 thousand records remained unlinked. The final data set contains 17,955 thousand persons.

Table 0.3. Final dataset from the combined PR/ER Register – Crime Suspects Register (CSR) linkage by type of linkage

	x 1,000
Total number of records in PR/ER/CSR file	17,955.0
Total number of CSR records in PR/ER/CSR file	1,969.0
Deterministically linked	1,487.0
To probabilistic linkage	482.0
Unlinked	408.8
Blocking 1 (postcode)	2.2
Blocking 2 (date of birth)	32.2
Blocking 3 (city + year of birth + month of birth)	14.1
Blocking 4 (city + year of birth)	16.4
Blocking 5 (city)	8.4

Data Quality Issues

Over 400 thousand records from the CSR remain unlinked. There are several reasons for this. First, the CSR contains all crime suspects. This also includes people who do not live in the Netherlands and can therefore not be expected to link to the PR or ER. This also includes people who no longer live in the Netherlands. There are also issues with the quality of the linkage variables. For a large number of records one or more variables used in the probabilistic linkage are missing. Therefore, even if these people are in the PR of ER, they cannot be linked because of missing information. Table 0.4 shows the percentage of missing values in the CSR for each of the linkage possibilities to the PR or ER. Of the records not linked to the PR or ER, almost 70% of the records

have a missing postal code or house number and 45% a missing city, house number or street. That makes is almost impossible to link these records as the only remaining linkage key is then the date of birth.

Table 0.4. Percentage of missing values in each of the variables from the CSR used in the probabilistic linkage after selection of target population.

PR	ER	No. records	Percentage of missing values in variable/variable combination						
			DOB	City	House numb.	Street	Postal code	Ci/Num/ Str	PC/ Num
N	N	7,780	0.0	42.5	40.8	44.2	68.5	46.2	68.5
N	Y	170	0.0	0.6	0.6	1.2	11.8	1.2	11.8
Y	N	7,710	0.0	1.0	1.0	1.1	1.9	1.1	1.9
Y	Y	2,330	0.0	0.8	0.8	0.8	1.5	0.9	1.5

Selection of target population

Before the capture-recapture method is applied the following operations are performed on the linked data set (PR/ER/CSR):

1. People can have several jobs, so in order to determine when a person was working in the Netherlands we merged subsequent job periods into one period. Gaps of 31 days and shorter are allowed. After the jobs have been merged, only persons present on 28th of March (this can be between two jobs) are selected from the ER. The merged job period is also used to determine the duration of work until the 28 March.
2. People in the CSR are only selected when they committed a crime in the first six months of 2014. However, for persons suspected of multiple crimes, only the date of the first crime, the date of the last crime and the number of crimes committed in 2014 were available. From this it is impossible to determine for all records whether or not a crime was committed in the first half of 2014. For people who committed multiple crimes in 2014 with a first crime before 2014 it is uncertain whether they have committed a crime in the first halve of 2014. For them it is assumed they were in the Netherlands in the first half of 2014.
3. From the PR we selected records that were registered in the PR on 28 March.
4. In order to derive the nationality of the persons the nationalities from the PR, ER and CSR are combined. With a preference given to first the PR, then the ER and finally the CSR. For those records without a known nationality the country of birth is used to derive the nationality.
5. Person from the Netherlands Antilles have the Dutch nationality. For the capture-recapture they are considered to have a separate nationality. Therefore, persons with a country of birth in the Netherlands Antilles except those whose parents are both born in the Netherlands, are recoded to a separate nationality 'Antillean'.
6. Nationality was recoded into 13 categories which were further recoded intro 7 categories (see table 0.5).

7. People whose country of residency is Belgium or Germany are removed from the dataset, as they are likely to live there and work in the Netherlands.
8. For records from the CSR, that are not linked to the ER or PR, no duration of their stay in the Netherlands is known. It is only known that they have committed a crime in the first six months of 2014. Some people in this group will have been less than twelve months in the Netherlands, so they do not belong to the target population. For these records the durations are imputed. A logistic regression model is estimated that predicts whether or not a person has been in the Netherlands for twelve months or more. This model is estimated using ER records that were not linked to the PR as these were considered to be most similar to the CSR records that were not linked to any register. The model uses age, gender and nationality (main nationality code) as predictors.
9. There is small number of records without age. For these records an age is imputed by selection random ages from the same nationality group (main nationality code).
10. Only records with an age between 15 and 65 are selected.
11. Only records in the PR or with a duration of stay longer or equal to one year are selected.

Table 0.5. Nationality codes used in the capture-recapture.

Main nationality codes		Nationality codes capture-recapture	
1	EU15 (excluding the Netherlands)	1	EU15 (excluding the Netherlands)
2	Polish	2	Polish
3	Other (New) EU	3	Other (New) EU
4	Balkan and former Soviet Union (excluding EU countries)	7	Other
5	Antillean and Surinam	5	Antillean, Surinam, Turkish and Moroccan
6	Turkish and Moroccan	5	Antillean, Surinam, Turkish and Moroccan
7	Iraqi, Iranian or Afghani	6	Asylum Countries
8	Other Middle East (excl. Israel) and North Africa	7	Other
9	South and South-East Asian and Middle and South America (incl. Mexico)	7	Other
10	East Asian (incl. China)	7	Other
11	Asylum Countries Africa	6	Asylum Countries
12	Other African (incl. South Africa)	7	Other
13	Other Western Countries	4	Other Western Countries

Annex 4 Estimation procedure of the unobserved usual residents

This annex discusses the method that underlies the estimate of the number of unobserved non-Dutch usual residents of age between 15 and 65. The data/registers that are used in this estimation procedure are discussed in the appendix on record linkage. The estimation procedure contains three methodological elements:

1. Imputation of missing data
2. Data selection
3. Sensitivity analysis
4. Capture-recapture (CRC) method

We will discuss these elements in this order.

Imputation of missing data
Imputation of missing data is important, because the applied estimation method can be sensitive to small groups. So for example, the records with gender and age 'unknown', can be a small group that leads to unreliable estimates.
There are two types of missing data:

1. Missing data on age and gender, which occurs in PR, ER and CSR
2. Missing data on duration of stay, which occurs only in the CSR.

Missing data on gender occurs in 0.16 per cent of the cases and missing data on age occurs in 0.10 per cent of the cases, measured over all ages. These missing values are replaced by imputed values, where the imputation is a drawing from the gender/age distribution of the nationality that belongs to the person represented in the record. Missing data on duration of stay can only occur in the CSR, because everyone in the PR is assumed to have a duration of over one year and all records in the ER have data on job duration, as explained in the appendix on record linkage. The missing data on the duration of stay is defined as a binomial variable with value 1 for a duration of over one year and 0 otherwise, corresponding to the definition of usual resident. The imputation method uses a log-linear model to estimate usual resident probabilities, where nationality, gender and age serve as explanatory variables, and the ER is used as the dataset to determine the model coefficients. These coefficients are used to calculate probabilities for records with missing values on the duration of stay in the CSR, which are used to sample values from a binomial distribution to replace the missing values.

Data selection
Data selection is important because the three registers need to cover the same target population, as this is a basic assumption that underlies the estimation method.
The three registers do not cover equal parts of the target population, i.e. usual residents between 15 and 65, so some data selection is required, i.e.:

- The PR covers all usual residents of all ages, so people under 15 and over 65 are removed from the PR, including the people with imputed values in that range.
- The ER contains people that stay in the Netherlands for less than one year, so these people, when they are not in the PR, are removed.

- The CSR contains people of over 65, so they are removed. Plus, people who, according to the CSR, stay in the Netherlands for less than one year (including imputed values on duration of stay) and are not in the PR, are removed from the data.

Sensitivity analysis

The quality of the CSR data is relatively poor, which may lead to unidentified erroneous captures or unidentified missed links. Gerritse et al. (2016) handles this issue by performing a sensitivity analysis, where different levels of erroneous captures and missed links are introduced, to see how they impact the estimation results. We replicated this approach and describe it here.

The sensitivity analysis is important because the CSR contains a large amount (68.5% of the records are unique in the CSR) of poorly linkable records (as described in the appendix on record linkage). This implies that these records are basically impossible to link (or to not link) to the PR and/or ER. Therefore records with little linkage information cannot simply be considered as unique captures that did not occur in the PR and/or ER, because that would lead to unrealistically high and biased population estimates. In order to see how large the impact of these potentially erroneous captures and missed links is, a number of scenarios is constructed, in which the records with insufficient linkage information are redefined as either erroneous captures or links with the PR and/or ER. This can be described as follows.

- A share of 68.5% of records in the CSR, that are not linked with the PR and ER and have no or little linkage information, are redefined as either erroneous capture (100%, 75%, 50%, 25% or 0%) or missed link (0%, 25%, 50%, 75% or 100%).
    - The missed links are redistributed as links with the PR and/or ER by the same distribution as the CSR records that are already linked to either the PR and/or ER.
- A small share (5%) of the remaining records (31.5%) in the CSR, could be erroneous captures or missed links due to the (poor) quality of the data. This smaller share is also redistributed as erroneous capture or missed link, in the same way as the records that contain no linkage information.

Additionally, a scenario is considered assuming that the CSR captures with no or limited linkage information are unique new captures that are not erroneous or missed links even though this scenario seems highly improbable. In total this would yield 21 scenarios (5 x 2 x 2 + 1). For each of these scenarios we can calculate a CRC estimate. In Gerritse et al. (2016) it is argued that not all scenarios are equally likely for the Dutch case. The main reasoning being that the police can make use of the PR when they register suspects in the CSR, whenever that person is present in the PR. Therefore, the records with incomplete linkage keys are very likely to be erroneous captures instead of missed links with the PR. Therefore, we will only show the results for the scenarios with 75% or 100% erroneous captures (and 25% or 0% missed links) along with the scenario where all records with limited linkage information are considered unique new captures (scenario 0).

Due to the random selection of erroneous captures and missed links, each of these scenarios contains a degree of randomness. In order to control for this randomness, we

repeat the calculation for each of the scenarios 20 times (a Monte Carlo approach) and use the average of these estimates as the CRC estimate.

Capture-recapture method

The CRC method has a long history going back to Petersen (1896), which will not be further discussed here. The CRC can be written as a Poisson regression model of a contingency table that contains registers and background variables as explanatory variables together with its corresponding frequencies. The Poisson model tries to explain the frequencies from the background variables, and uses these relations to 'predict' the number of records that are not captured in any register. In line with Gerritse et al. (2016) the CRC is performed separately for seven nationality/region levels. The seven groups of nationalities that are distinguished, are composed mainly on their geographical proximity, general size and relevance to the Netherlands. These groups are described as: 'EU15', 'Polish', 'Other EU', 'Balkan and other former Soviet states', 'Turkey, Morocco and others', 'Iraq and others' and 'Other western countries'. Furthermore, the background variables gender and age serve, together with the three registers, as explanatory variables. The model also takes into account all the first- and second-order interaction effects between the explanatory variables, except for the second order-interaction between the three registers, which ensures the model is identified. The model is estimated in R with the function *glm* with *family = poisson()*. Furthermore, in order to select the most parsimonious model according to the Bayesian Information Criterion (BIC) (as in Gerritse et al. (2016)), the *step* function is used, which applies a general-to-specific approach.

The CRC estimates for unobserved usual residents for the base scenario 0 and the more likely scenarios with relatively large portions of erroneous captures, are summarised as follows:

Scenario	CRC estimate (x1000)	Monte Carlo confidence interval (x1000)	Share of erroneous captures among the CSR unlinkables	Share of missed links among the CSR unlinkables	Share of erroneous captures among the CSR linkables	Share of missed links among the CSR unlinkables
0	506	(506-506)	0%	0%	0%	0%
1	124	(122-126)	75%	25%	0%	0%
2	125	(123-127)	75%	25%	5%	0%
3	116	(113-119)	75%	25%	0%	5%
4	117	(115-119)	75%	25%	5%	5%
5	136	(136-136)	100%	0%	0%	0%
6	136	(136-136)	100%	0%	5%	0%
7	126	(122-130)	100%	0%	0%	5%
8	125	(122-128)	100%	0%	5%	5%

The first thing that is apparent from this table, is that the estimate of the first scenario is far higher than the other estimates. This is simply the result of assuming that the CSR captures with no or limited captures are unique new captures that are not erroneous or missed links, which seems highly improbable. The other scenarios show much lower estimates together with tight confidence intervals. Moreover, these estimates are all

relatively close to each other, which implies that it hardly matters how captures with no or limited linkage information are redefined as erroneous capture or missed link. Furthermore, we see that redefining a small share of unique CSR captures with sufficient link information (column five and six) also has little impact.
Using the confidence intervals found by the Monte Carlo simulations and the variation over the scenarios 1-8, we find that the estimate of the unobserved usual residents between 15 and 65, by either the PR, ER and CSR would be between 113 and 136 thousand.

Annex 5 Usual residents metadata forwarded to Eurostat[6]

Concept name	DEMOMIGR_UREESMS_A_NL_2013_0000
1. Contact	
1.1. Contact organisation	Statistics Netherlands
1.2. Contact organisation unit	Demographic statistics
1.3. Contact name	Wim Leunis
1.4. Contact person function	Teammanager
1.5. Contact mail address	P.O. Box 24500 | 2490 HA The Hague
1.6. Contact email address	demographicsNL@cbs.nl
1.7. Contact phone number	+31(0)70 337 5474
1.8. Contact fax number	
2. Metadata update	Not applicable
2.1. Metadata last certified	Not applicable
2.2. Metadata last posted	Not applicable
2.3. Metadata last update	Not applicable
3. Statistical presentation	

[6] 2016 meta data were on line forwarded to Eurostat by ultimo August 2016.

3.1. Data description	Total usual residence population for the purposes of qualified majority voting in the Council.
3.2. Classification system	Optional
3.3. Coverage - sector	Total usual residence population at national level.
3.4. Statistical concepts and definitions	Those who have lived or intend to live for a period of more than 12 months in their place of usual residence.
3.5. Statistical unit	Person.
3.6. Statistical population	Optional
3.7. Reference area	Data are available at national level.
3.8. Coverage - Time	Optional
3.9. Base period	Optional
4. Unit of measure	Person.
5. Reference Period	The reference date for population data is the end of the reference period (midnight of 31 December).
6. Institutional Mandate	
6.1. Institutional Mandate - legal acts and other agreements	Optional
6.2. Institutional Mandate - data sharing	Optional
7. Confidentiality	
7.1. Confidentiality - policy	Optional

7.2. Confidentiality - data treatment	Optional
8. Release policy	
8.1. Release calendar	Optional
8.2. Release calendar access	Optional
8.3. Release policy - user access	Optional
9. Frequency of dissemination	Annual.
10. Dissemination format	
10.1. Dissemination format - News release	Optional
10.2. Dissemination format - Publications	Optional
10.3. Dissemination format - online database	Optional
10.4. Dissemination format - microdata access	Optional
10.5. Dissemination format - other	Optional
11. Accessibility of documentation	

11.1. Documentation on methodology	As far as the options mentioned by Eurostat are concerned, the Netherlands use an option B method (Total national population definition has been adjusted to the total usual definition) To estimate the population of usual residents, capture – recapture methods are available. By making a three list estimation, making restrictions to one day for the period based registers and a short period for the event based register, applying a stringent linkage method and deletion of erroneous records, most of the assumptions of the capture – recapture method are met. Calculations have been made for ultimo September 2010. Under the assumption that the relative difference between register and usual residence population had not been changed between this date and 01-01-2014 (not likely, but information lacks to present a better one) this figure has been adjusted with the registered population change between 2010 en 2014. To estimate the number of usual residents, the estimated total number of persons has been divided into those who stay longer than a year in The Netherlands and those who do not. Here a distinction is made between the category already registered in the population register who are assumed to meet the 12 months criterion and the residence duration from employment records by making use of the starting and ending dates of sequential jobs in the employment register. When there are gaps between jobs, the sensitivity of the effect of different durations of these gaps have been analysed. For the third register used imputations had te be made on duration of residence.
11.2. Quality management - documentation	Optional
12. Quality management	**Quality management** The estimates are provisional and temporary. A more detailed description of the applied methods could be found in Van der Heijden et al. (2012), Gerritse et al. (2015), Bakker et al. (2014). **Capture-recapture method** To estimate the total number of persons, whether they are registered or not, capture – recapture methods are available. Two or more registers that contain information on (parts of) the population are linked and subsequently the missed part of the population that have not been captured in either list is estimated. A two list version of this method has the following assumptions: (a) the population should be closed during the data collection period; (b) the inclusion probability of being registered in the first list is independent of the inclusion probability of being registered in the second list; (c) the elements of the population should have a positive probability of being registered in the registers; (d) the lists are perfectly linked; (e) the sources do not include erroneous captures. Violation of these

	assumptions could lead to highly biased estimates. We took the following precautions: - To relax the assumption of independency between the registers, we make use of three registers: Dutch Population Register (PR), the Employee Register (ER) and Crime Suspects Register (CSR). Because the ER contains only employed persons, the age range is between 15 and 65. The young and elderly are estimated in another way. - The assumption that the population is closed, is satisfied for the PR and ER by restricting the data to the status of only one day. This cannot be applied to the CSR because this register is event based: crimes are registered of which the police make a report. In order to satisfy the assumption as good as possible, we restrict the period of the CSR to the second half of 2010. - To prevent erroneous captures, one needs to remove the records of persons from the three lists who do not belong to the population. Because it is not known which records do belong to the population or not, we simulated this by assuming that 10, 20 or 30% of the captures in the CSR that did not link to the PR and ER are erroneous captures. In that way, we produced a range of estimates. - The capture –recapture method is very sensitive to linkage error. Therefore, the three lists are linked pairwise with much caution. Despite the attention paid to the linkage method and the careful execution of the method, it is still possible that not all records are linked that should be linked. The records in the CSR that do not link to either the PR nor the ER, contain relatively large numbers of missing values in the linkage key. Therefore, it is to be expected that for that reason a large number of missed links occur. We used this information to produce a range of estimates by assuming that 10% above the 30% erroneous captures are missed links. Because we assume that there is a lot of overlap between missed links and erroneous captures, we restrict it to 10% extra. The number of children aged 0-15 is estimated by using education registers and is estimated on 3600 usual residents. The number of elderly (aged 65 and older) has been estimated by extrapolating the results for the age class 55-64 years. It is estimated by approximately 2 to 10 thousand (Bakker et al., forthcoming) The analyses result in a range of estimates. The lowest estimation for October 2010 is an undercount of 190 thousand, the highest one an undercount of 311 thousand. According to the Eurostat regulation we have to present one figure as a result. In order to do this, we took the mean of the minimum and maximum estimation of the undercount, because information lacks to decide for the most probable figure: 251 thousand.

Taking the development between 01-10-2010 and 01-01-2014 into account, the undercount for 01-01-2014 varies between 192 en 215 thousand, with a mean of 253 thousand, which leads to an estimated number of usual residents of 17.082 thousand

References

Bakker, B.F.M., S. C. Gerritse, P.G. M. van der Heijden, J. van der Laan, R. van der Vliet, M. Cruyff, 2014, Estimation of Usual Residents in Register Based Statistics in the Netherlands (forthcoming)

Gerritse, S., P.G.M. van der Heijden & B.F.M. Bakker, 2015, Sensitivity of population size estimation for violating parametric assumptions in loglinear models, accepted for publication by Journal of Official Statistics

Heijden, P.G.M. van der, J. Whittaker, M. Cruyff, B..Bakker & R. van der Vliet, People born in the Middle East but residing in the Netherlands: invariant population size estimates and the role of active and passive covariates, *Annals of Applied Statistics,* vol. 6, nr. 3, pp. 831-852

12.1. Quality assurance

12.2. Quality management - assessment

12.2.1 Nationals

Please indicate the type of duration of stay when counting a person with national citizenship as an usual resident with a minimum duration of stay of 12 months according to Art. 2(d) of Reg. 1260/2013.

1. Actual	2. Intended	3. Actual and intended	4. Other. Please specify below:
X			

12.2.2 EEA Citizens

Please indicate the type of duration of stay when counting a person with EEA citizenship as an usual resident with a minimum duration of stay of 12 months according to Art. 2(d) of Reg. 1260/2013.

1. Actual	2. Intended	3. Actual and intended	4. Other. Please specify below:
X			

	12.2.3 NON-EEA Citizens Please indicate the type of duration of stay when counting a person with non-EEA foreign citizenship as an usual resident with a minimum duration of stay of 12 months according to Art. 2(d) of Reg. 1260/2013. 1. Actual: X 2. Intended: 3. Actual and intended: 4. Other. Please specify below:
13. Relevance	
13.1. Relevance - User Needs	Optional Data requested in Art. 4.1 of Reg. 1260/2013.
13.2. Relevance - User Satisfaction	Optional
13.3. Completeness	Optional Data transmitted as requested in Art. 4.1 of Reg. 1260/2013.
14. Accuracy	
14.1. Accuracy - overall	
14.2. Sampling error	Optional
14.3. Non-sampling error	Optional
15. Timeliness and punctuality	
15.1. Timeliness	Optional For the purposes of qualified majority voting in the Council, Member States shall provide the Commission (Eurostat) with data on the total population at national level at the reference time, in accordance with Article 2 (c), within eight months of the end of the reference year.
15.2. Punctuality	Optional
16. Comparability	
16.1. Comparability –	Optional

geographical	
16.2. Comparability - over time	Optional
17. Coherence	
17.1. Coherence - cross domain	Optional The total usual residence population transmitted by the national statistical offices to Eurostat in the framework of the Usual Residence Population data collection may differ from those available in National Accounts, Labour Force Survey or in the survey on Income and Living Conditions. The difference is given by the population coverage (see metadata specific to each domain).
17.2. Coherence - internal	Optional The total usual residence population transmitted by the national statistical offices to Eurostat in the framework of the Usual Residence Population data collection may differ from those available in the rest of the Demographic domain. The difference is given by the population definition (see metadata specific to each domain).
18. Cost and Burden	
19. Data revision	
19.1. Data revision - policy	Not applicable.
19.2. Data revision - practice	Not applicable.
20. Statistical processing	
20.1. Source data	1. Dutch population register 2. Employee register 3. Registered crime suspects
20.2. Frequency of data collection	Annual.
20.3. Data collection	
20.4. Data validation	The total population undercount according to this model is for October 2010 estimated between 190 to 311 thousand persons who are not registered in the PR. This is more or less in line with previous estimations as

	Hoogteijling (2002) and Bakker (2009) report. The first one estimated the non-registered for 2001 between 46 and 116 thousand. The second one estimated the number of non-registered at 236,5 thousand in 2006. We should expect that the number of not-registered persons increased since then, because of the changed position of the inhabitants of new EU-countries. This has led to a large migration flow to the Netherlands from, in particular, Polish and other eastern European countries. *References* Bakker, B.F.M., 2009, Trek alle registers open! (Amsterdam: VU University Press) Hoogteijling, E.M.J., 2002, Raming van het aantal niet in de GBA geregistreerden, Rapport 177-02-SOO (Voorburg: Centraal Bureau voor de Statistiek) http://www.cbs.nl/NR/rdonlyres/38D3F9AB-A9C7-4D99-92F8-C8AD700D792/0/ramingnietGBAgeregistreerden.pdf
20.5. Data compilation	Not applicable.
20.6. Adjustment	
21. Comment	